Property Model-based Tailor-made Design of Chemical-based Products

Kalakul, Sawitree

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Property Model-based Tailor-made Design of Chemical-based Products

Sawitree Kalakul

PhD Thesis

May 2016

DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Property Model-based Tailor-made Design of Chemical-based Products

PhD Thesis

Sawitree Kalakul

May, 2016

PSE for SPEED Project
Department of Chemical & Biochemical Engineering
Technical University of Denmark

PREFACE

This thesis is submitted as partial fulfilment of the requirements for the degree of Doctor of Philosophy (PhD) at the Technical University of Denmark (DTU). This work has been carried out at SPEED Project of the Department of Chemical and Biochemical Engineering, from June 2013 to May 2016, under the supervision of Professor Rafiqul Gani and Professor Georgios M. Kontogeorgis, from the Center for Energy Resources Engineering (CERE). The financial support for this PhD project was funded by a NPRP award [NPRP 05-066-2-023] from the Qatar National Research Fund (a member of the Qatar Foundation), Alfa Laval Copenhagen and DTU.

Leaving my country (Thailand) for doing PhD in Denmark and departing from +36 to -6 degree including culture differences were very challenging for me. Nonetheless, these obstacles could not be overcome without the following unforgettable persons.
I am grateful to my supervisors, Professor Rafiqul Gani and Professor Georgios M. Kontogeorgis for their valuable input, guidance and motivation throughout this project. My special thanks go to my main supervisor, Professor Rafiqul Gani for his patient guidance, enthusiastic encouragement and having faith in me by giving me opportunities to work in the awesome projects. I have learned a lot from you, more than I expected during my PhD.

I would also like to extend my thanks to my friends/colleagues in the SPEED Project for their moral and/or technical supports. Last but not the least, I would like to express my gratefulness to my family; mom and dad and all of my friends in Thailand and Denmark, especially, Pratoom Jindakum, Siriwan Thansupapol, and Thomas Malthe Trier Jacobsen for all of your support throughout my PhD study.

ABSTRACT

Computer-aided model-based methods and tools are increasingly playing important roles in chemical product design. They have the potential to very quickly search for and identify reliable product candidates that can then be verified through experiments. In this way, the time and resources spent on experiment are reduced leading to faster and cheaper to market the products. The tools also help to manage the solution of product design problems, which usually require efficient handling of model-data-knowledge from different sources and at different time and size scales.

The main contribution of this project is: (1) the development of a systematic model-based framework for chemical product design; (2) its implementation as a computer-aided tool based on a specially developed architecture; (3) the creation of product design template together with their algorithms, models, tools and data for various types of products. The goal has been to develop a chemical product simulator, similar in concept to a process simulator, which make the product design and development easier and faster, and provide the way for unified and consistent product documentation. In the same way a typical process simulator works, the developed product simulator (VPPD-Lab) allows product designers to; (1) analyze chemicals based products by performing virtual experiments (product property and performance calculations); (2) predict the properties of products; and (3) create new product property and product performance models, when needed. However, unlike process simulators, VPPD-Lab can also be used directly for (4) design of chemicals based products using the design template for various types of products, such as, single molecule products, formulations, blends, emulsions and devices; and, (5) creation of new product design templates when the needed template for a desired product is not available. VPPD-Lab employs a suite of algorithms (such as database search, molecular and mixture blend design) and toolboxes (such as property calculations and property model consistency tests) for specific product property prediction, design, and/or analysis tasks.

In order to achieve the features mentioned above, several issues need to be addressed: the translation of consumer needs into target properties; property models and available data for each type of chemical products; design methods and algorithms; available computer-aided tools; the systematic framework for chemical product design and analysis and its implementation as architecture for VPPD-Lab. From many test problems, eight application examples are presented to illustrate the use of the software. For two of these examples, the prediction of product properties and the use of virtual experiments to test product performances are highlighted. Five examples illustrate the use of the product design templates with respect to five types of chemical products (molecular design, formulation design, emulsion design, blend design and device design).

RESUME PÅ DANSK

Computerstøttede modelbaserede metoder og redskaber spiller en vigtigere og stigende rolle i kemisk produktdesign. De har potentiale til meget hurtigt og pålideligt at finde produktkandidater, som kan verificeres gennem forsøg. På den måde vil den tid og de ressourcer, som bliver brugt på forsøg blive reduceret, hvilket vil gøre det hurtigere og billigere at markedsføre produkterne. Redskaberne hjælper også til at styre og løse problemer med produktdesign, hvilket ofte vil kræve en viden om modeldata fra forskellige kilder og forskellige tids- og størrelsesskalaer, for at det kan håndteres effektivt.
Det væsentligste bidrag for dette projekt er 1, udvikling af systematiske modelbaserede rammer for kemisk produktdesign; 2, dets implementering af computerstøttede redskaber baseret på specielt udviklet arkitektur; 3, skabelsen af en produkt designskabelon sammen med deres algoritmer, modeller, redskaber og data for forskellige typer af produkter. Målet var at udvikle en kemisk produktsimulator lignende en processimulator, hvilket gør produktdesign og udvikling nemmere, hurtigere og giver en forenet og konsistent produktdokumentation. På den samme måde som en typisk processimulator virker, vil den udviklede produktsimulator (VPP-Lab) tillade produkt designere at 1) analysere kemisk baserede produkter ved at udføre virtuelle forsøg (produkt egenskaber og udførelse af beregninger); 2) forudsige produktegenskaberne; og 3) skabe nye produktegenskaber og produktudførelsesmodeller, når behovet opstår. Dog i modsætning til proces simulatorer kan VPPD-Lab bruges direkte til 4) design af kemisk baserede produkter ved brug af designskabelonen til forskellige typer af produkter som enkelt molekylærprodukter, formuleringer, blandinger, emulsioner og enheder; og 5) skabelse af nye produkt designskabeloner, når den nødvendige skabelon for et ønsket produkt ikke er tilgængelig. VPPD-Lab benytter en række algoritmer så som databasesøgninger, molekylær- og miksturblandingsdesign og redskabskasser så som beregninger af egenskaber og modelegenskabernes konsistenstests til forudsigelse af specifikke produktegenskaber, design og/eller analyseopgaver.

For at opnå de ovenfor nævnte funktioner må flere spørgsmål besvares: Kendskab til brugerens behov for slutegenskaber; modelegenskaber og tilgængelig data for hver type af de kemiske produkter; design metoder og algoritmer; tilgængelige computerstøttede redskaber; de systematiske rammer for det kemiske produktdesign og analyser samt dets implementering som en arkitektur i VPPD-Lab. Otte applikationseksempler er blevet udviklet for at illustrere brugen af softwaren. For to af disse eksempler er forudsigelsen af produktegenskaberne det virtuelle eksperiment på produkt udførelsen fremhævet. Fem eksempler illustrere brugen af produktdesignskabelonen med hensyn til fem typer af kemiske produkter (molykulærdesign, formuleringsdesign, emulsionsdesign, blandingsdesign og enhedsdesign).

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

The chemical industry is a vast industry with a wide variety of chemical products (CPs), including agrochemicals, ceramics, elastomers, electronic materials, explosives, foods, flavors and fragrances, fuels, industrial gases, inorganic chemicals, metals, oleochemicals, petrochemicals, pharmaceuticals, plastics, and textiles. Due to the variety of these CPs, the topic of chemical product design and analysis (CPD) is attracting increasing attention.

1.1 Chemical product design and analysis (CPD)

Nowadays chemical industries are increasingly moving beyond the production of commodity chemicals (characterized by large quantity, continuous production of low added-value molecules) to specialty chemicals (characterized by small quantity, batch production of high added value) as well as formulated products (characterized by complex mixtures targeted to match desired end-use properties) (Hill, 2009). Due to the changing nature of the chemicals based industrial sector, chemical companies are having to consider new technologies as well as handle marketing challenges (time to market, research and development costs and efforts for developing new smart products, consumer satisfactions). While the design of commodity chemicals (such as ammonia, sulphuric acid and methanol) only focuses on process optimization in order to maximize profits due to their uncomplicated molecular structures, the design of CPs (such as paint, lotions, cream, foods, drug solutions and polymers) are more complex and involve a variety of additional issues not commonly encountered in process design (Cussler and Moggridge, 2000). While process design and optimization are fundamental to chemical engineers, product design should also play an important role because a process is needed only when a selected product needs to be made. Designing a promising product, may in some cases need relevant contributions from chemistry, mathematics, modeling, material sciences, thermodynamics, kinetics, and many more topics. Cusslerr and Moggridge (2003) include process design and manufacturing as one step of product design and also emphasize the importance of CPD in chemical engineering.

The goal of chemical product design is to find a product that exhibits a set of desirable or specified behavior. The chemical product design problems involve establishing a list of candidates satisfying given targets and choosing the most appropriate candidate to be verified by experiments. **Figure 1.1** represents a chemical product tree that gives an idea of the variety of chemical products that can be derived from raw materials (simple products); as one moves higher up the tree, more complex structured products with higher added value are obtained (Gani, 2004). The root of the tree consists of a number

of raw materials (approximately 10) which are processed to obtain the commodity chemicals. Intermediate chemicals are then manufactured from the commodities. Finally, the leafs of the tree represent a very wide range of high added value products (almost 30,000 products) obtained by processing and/or combining the lower level products (raw materials, basic and intermediate products).

Figure 1.1 Chemical product tree and classification of chemical-based products (Gani, 2004)

Chemical engineers are involved in the successful development, application and/or manufacturing of these products. In order to design these products, Cussler and Moggridge (2003) propose a general four-step approach:

(1) Identification of consumer needs that should be met by the product;
(2) Generation of ideas that can satisfy the needs;
(3) Selection of the most promising product idea;
(4) Development of a process to manufacture the desired product.

This approach is applicable to all types of CPs but the work-flow for specific products are different because different products involve different issues, needs and therefore, require different aspects of science and technology. While the performances of single molecule products are based on their molecular structures, the refined chemicals and consumer products are related not only to the presence of active ingredients and additives in the formulation but also to the product structural and material properties (Smith and Ierapepritou, 2011). For example, food packaging has 3 main layers (outside print layer, adhesive layer and inside barrier layer); each layer is covered by films made of different materials. While it is not stated how these steps should be performed, for steps 2 and 3 use of experiment-based trial and error approaches is quite common and for step 4, which is regarded as the process design problem, use of a wide range of methods and tools is already possible now (Cussler and Moggridge, 2000). Usually, the experimental-based tasks are expensive and time-consuming. Gubbins and Quirke (1996) highlight the cost issue related to experimental tests and suggested a cost of 2600 US$ per one data point for the design of a molecule. In experiment-based approaches, a set of experiments are usually performed and the results are compared in order to select the best solution that might or might not be regarded as the optimal solution. However, competitiveness in the market is forcing companies to produce better products and look for a "first time right" production or even the adoption of new aims such as "one customer = one product" (Harper, 2008). Therefore, it is now generally accepted that application of model-based approach helps to design/improve products to reach the market faster by reducing some costly and time-consuming experiments (Gani, 2004). That is, mathematical models derived by translation of product behavior to equivalent physico-chemical properties are solved and numerical results are interpreted analyzed for verification of product end-use properties and/or product behavior. Through computer-aided tools, property models are used to generate product data corresponding to thousands of design ideas (as input to the models). This way, product candidates are generated and screened in order to obtain the optimal product, which then is validated by means of dedicated experiments (Charpentier, 2009).

Mathematical modeling has been introduced (Achenie and Wang, 2002), especially in applications where product performances and properties are related to the molecular structures of chemicals. It forms the basis of a hybrid global optimization approach for solving solvent design problems modeled by mixed integer nonlinear programming (MINLP). Promising work in computer aided molecular design (CAMD) has been summarized in the book by Achenie, Gani, & Venkatasubramanian (2002). CAMD has been applied to design various single molecule products such as solvents for separation (Hostrup et al., 1999; Chemmangattuvalappil et al., 2010), solvents for organic synthesis (Gani et al., 2005), design polymers (Satyanarayana et al., 2009), refrigerants, active ingredients and many more. Computer-aided mixture/blend design (CAM^{b}D) has been applied for solvent mixture design (Modarresi et al., 2008; Eden et al., 2004). Yunus et al. (2014) employs decomposition approach to formulate and solve tailor-made blended design problems. However, capturing the knowledge from product designers and experts and storing it in terms of explicit knowledge enables the use for the design and development of innovative products (Joglekar et al., 2014). This approach has been used for design of homogenous formulated product such as, an insect repellent lotion and paints (Conte et al., 2011), as well as emulsified products, such as, hand-wash detergent and sunscreen lotion (Mattei et al., 2014).

Within process system engineering (PSE), CPD is a comparatively new field compared to process design, which has reached a high degree of scientific maturity (Bagajewicz, 2007). Therefore, CPD has been identified as a challenge as well as means of great potential benefits for chemical industries. The limitation of computer-aided techniques is related to the limitation of property models as well as the complexity of the systems as highlighted by Gani (2004). Furthermore, Gani (2004), Gani and Ng (2015) highlight the need for a multidisciplinary approach to have more control in the end-use characteristics of chemical products. Thus, product design/engineering proposed as the third paradigm (Voncken et al., 2004; Cussler and Wei, 2003) should be interpreted to include computer-aided approaches (making use of new modeling and simulating tools) including the multiscale and multidisciplinary modeling approach in order to perform design of the product at different levels of abstraction and observation. All these modeling needs, incorporated in one appropriate knowledge-based and model-based library, together with different data-flow and work-flow templates (for design of different products), could be accessed through a framework for chemical product-process design. This is the main driving force for the developments reported in this PhD-thesis.

1.2 Motivation and objectives

The main objective of this PhD project is to extend the application range of CPD. As highlighted in the discussion above related to CPD, some challenges in the area of computer-aided model-based approaches have been identified. Note that, the main goal in chemical product design is to design a final product with the required end-use characteristics desired by the consumer. Huge amounts of data/models, methodologies and algorithms are being developed and research involving the development of new methodologies to estimate target product properties are being conducted. The key that allows wider application ranges for solving each type of product design problem is to store the corresponding product design work-flows, data-flows, tools, models and calculation algorithms within a framework that is easy to maintain, update and apply. The framework needs to incorporate templates for CPD problems. The template idea is a way to accommodate the different needs (properties, models, work-flows, date-flow, etc) for different classes of chemicals based products. Therefore, the framework together with the template is able to integrate computer-aided tools so that it is able to perform product design and analysis involving wide ranges of CPs considering multiscale and multidisciplinary modeling in a systematic manner. The advantage of having a framework is that it would serve as the glue that puts everything together. It should also be possible for the framework to capture past experiences in order to provide better guidelines for future CPs (Gani, 2004). Thus, the main objectives of this PhD project are the following:

- Development of template-based approach for chemical product design, which emphasizes the use of a product design methodology, property models and product design knowledge not only for one specific application but also for future applications that have the same product nature and disciplines with the parent template.

- Creation of a systematic framework for chemical product synthesis, design and evaluation which should include multiscale and multidisciplinary features that covers a very wide range of chemical-based products to provide structures/ work-flows, supports and guidance to solve the current and future product design problems. This is a challenging task requiring data acquisition, data testing, model development and multi-scale modeling that needs to be integrated within a product design framework.
- Implementation of the framework that serves as a software architecture, into a computer-aided model-based tool (VPPD-Lab) that can be employed for solution of a wide range of design, analysis, evaluation problems in a fast, efficient, reliable and systematic manner.
- Testing and validation of the VPPD-Lab options with various product design case studies: tailor made design of jet-fuels such as gasoline and diesel; formulated products such as insect repellent, hair-spray and sunscreen lotions; emulsified products such as a hand-wash detergent and a cleaning detergent; single molecule designs such as solvents and refrigerants; and devices such as micro-capsule for a controlled release of a pesticide.

1.3 Structure of the thesis

The contents of this PhD-thesis are divided into six chapters, including this current chapter (introduction); where, a brief overview of the work within the context of chemical product design and analysis is given. Chapter 2 is concerned with CPD background including the classification of CPs, activities involved in CPD and CPD approaches. Some of the important issues related to product design and analysis are discussed. Chapter 3 introduces the concept of product design template made available through a framework that includes a collection of methodologies, design algorithms, databases, property models, etc., needed for CPD. Chapter 4 presents the implementation of the framework in the software called VPPD-Lab. Chapter 5 presents the results from various CPD related case studies highlighting different applications of the software. Chapter 6 summarizes the main achievements of this PhD project and gives an overview of some of the challenges and issues that need to be addressed in future.

2 PRODUCT DESIGN: BACKGROUND

CPD covers an enormously wide area and a wide variety of activities are involved in it. This chapter introduces the tasks and phases in CPD and development, the principal issues and needs for different classes of product design problems and the challenges and opportunities for the PSE/CAPE community with respect to developing systematic methods and tools that can contribute positively towards their solution.

2.1 Classification of chemical products

CPs can be classified in various ways. In this thesis, CPs are classified into five types (single molecular products, blended products, homogeneous formulated products (formulations), emulsified products (emulsions) and devices, which are adapted from Ng and Gani (2015). In the text below, each product type is briefly explained:

- Single molecular products – Single molecule products can be found in many market sectors: pesticides in agricultural production; sugar ester in food and beverage production; and solvents in medicine production processes. They are obtained from the processing of the raw materials (oil, gas, etc.) in very large quantities. They are sold on the basis of their purity;
- Blended products – These products refer to the formulation of various single molecules in a single-phase blended liquid product (such as synthetic fuels and lubricants) or homogenous polymer blends. These products often have the main ingredients (MIs) that perform the main functions of the product and additives that enhance their qualities;
- Formulations (single-phase products) – These products are commonly found in cosmetics and food consumer goods. They are constituted of several ingredients that are combined together into one phase (usually liquid). For example an insect repellent lotion contains a solid active ingredient (AI) that is responsible for the main function, solvents that dissolve and deliver the AI, and additives that enhance its quality;
- Formulations (emulsified products) – Emulsions are defined as dispersed systems for which the phases are immiscible or partially miscible liquids. Emulsions are dispersions of one liquid in another, such as oil-in-water (cream, detergents) or water-in-oil (butter) emulsions. They are characterized by droplet size of about 1 μm, and they are typically unstable systems, which will

eventually separate and require emulsifiers, most often surfactants, to be kinetically stabilized. They are found in the food, pharmaceutical, and cosmetic industries.

- Devices – These products are used to measure, make, purify, or transform or transfer chemicals (Seider et al., 2008). Chemical devices are often like small chemical plants, taking a chemical feed stock and producing a product. For example: a humidifier boils water and releases the warm steam into the room to increase the humidity; and a catalytic converter is used to convert harmful pollutants into less harmful emissions before they leave the car exhaust system. Chemical devices are often like small chemical plants, taking a chemical feed stock and producing a product.These products involve in phenomena such as reactions, fluid flow, heating, cooling and separations.

2.2 Activity diagram for CPD

Product design and development involve a wide variety of issues. The starting point of a product design and development project is the creation of an objective-time chart (Ng, 2003) where a period of time is specified in order to perform and archive each objective of product design and analysis. Cheng et al (2009) propose a rule-based base method called RAT^{2}IO module to assist activities related to each objective in an objective-time chart as shown in **Figure 2.1**. The acronym stands for Resources, Activities, Time, Tools, Input/output information, and Objectives. Thus, resources (such as money and people) are required in order to perform certain activities (such as experiments and modeling) within the specific period of time by means of tools (such as computer software and experimental setup) to generate outputs that meet each objective. Cheng et al (2009) list seven main objectives related to chemical product design and development:

(1) Project management;
(2) Market survey;
(3) Product conceptualization;
(4) Design product formula;
(5) Manufacturing planning;
(6) Financial analysis;
(7) Market testing.

Each objective involves different input information, activities, tools, resources and time. In project management, business goals are the input information for project managers to set objective time chart and checking the project progress. Marketing teams make surveys to find consumer needs and analyze the survey results to set preliminary price ranges, a project image and selling points. In product conceptualization, consumer needs to specify technical specifications. Product images and selling points are used to identify the product microstructure. Chemical engineers are mainly involved in objectives (4), (5) and (6) of the overall project. In design product formula, product engineers translate the technical specifications into physical and chemical properties and try to find a prototype formula with refined concentration of active ingredients that

satisfies the physical and chemical properties by means of experimental tests. In manufacturing planning, process engineers together with environmental engineers are involved in manufacturing process alternatives, waste treatment planning with respect to regulations on environmental issues and government regulations for plants and processes. In financial analysis, process operating cost, capital cost, product selling price, financial return are estimated by financial controller and chemical engineers based on the predefined equipment, raw materials and process alternatives. Finally, marketing and testing can be done by the production teams together with the marketing team and sales teams to carry out pilot scale production; distribution and testing of samples to potential buyers; and development of marketing plans in order to identify key buyers and sale channels.

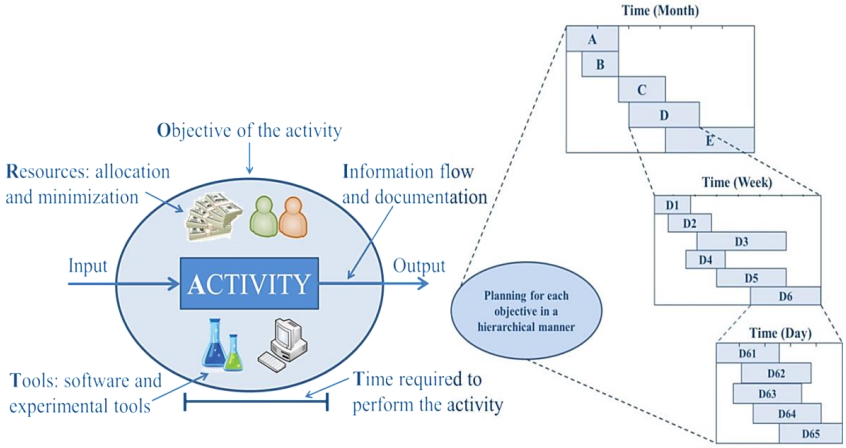


Figure 2.1. A generic objective-time chart and RAT^{2}IO module (Cheng et al., 2009)

In general, the most time consuming objective is the design product formula (more than 30 % of the overall project time is spent) which is longer than manufacturing planning together with financial analysis (Cheng et al., 2009). Therefore, nowadays, many researchers focus on designing CPs using different solution approaches as described below.

2.3 CPD approaches

The CPD approaches can be classified into the following types (Ng, Gani and Dam-Johansen, 2007):

- Experimental-based trial and error approach – This approach has been employed widely for the design and development of CPs for decades. The objective of this approach is to characterize the properties of the product candidate as well as the product prototype; to verify if these properties match the requirements of the product and to change the product composition (formula) until the requirements are satisfied. This approach needs the use of knowledge of experts in terms of

heuristics and guideline in order to generate lists of product candidates, adjust the product compositions to target values. This approach is employed when mathematical models for the estimation of product target properties/functions are not available. Past knowledge and experience of product designers are crucial in this approach. Since the desired product performances or functions need to be measured, the disadvantages of this approach are: long development times, high consumption of resources, need of the experts on specific type of products. Therefore, this approach is employed when there are not many candidates or the end-use properties need to be verified before manufacturing.

- Model-based approach – This approach is used when validated mathematical models for the estimation of target properties are available. This approach aims at screening numerous product alternatives in order to identify a smaller number of candidates by means of computer-aided tools. However, the uncertainty and application range of this approach are based on the availability and reliability of property models. For example, some target properties (such as scent or appearance) cannot be modeled.
- Integrated experimental-modeling approach – This approach is used when mathematical models are not available for all target properties. Predictive models are employed to generate and test numerous candidates and identify a small number of promising candidates that will further be investigated through more rigorous models, correlations and/or experiments. That is, the uncertainties of the model-based approach are compensated by the experimental part while the number of experiments is reduced through the model-based approach. Therefore, the search space is reduced and time and resources can be spared. The expensive experiments are reserved only for the most promising candidates.

The proposed product design methodologies employed in this PhD project are based on the integrated experiment-modeling approach.

2.4 Computer-aided product design

The CPD problem is formulated as: given a set of desired specifications (such as physio-chemical properties for a product), determine the chemical product that satisfies the *a priori* defined targets. This problem is usually described as the 'reverse property prediction problem' (Gani and Pistikopoulos, 2002). In property prediction problems, the chemical structure of the compound is known and the properties are calculated through property models. In a product design problem, the desired compound properties are known and the chemical structure of the compound needs to be identified. Most of product design problems employ 'Generate and test' algorithm where the property prediction is solved repeatedly to test the generated alternatives. However, chemical products are so diverse, that it is very difficult to develop a general methodology for all kinds of products (Gani and Ng, 2015). Therefore, a wide variety of computer-aided methodologies and tools have been developed based on types of chemicals based products. These methods can be presented as following:

Computer-aided molecular design (CAMD)

This methodology is used to design a pure compound (a single molecule product). The design problem is defined as follows: given a set of building blocks and a specified set of target properties, determine the molecule or molecular structure that matches these properties (Gani, 2004). The molecular structure of a compound is usually represented through groups (Harper et al, 2000) and/or connectivity indices (Camarda and Maranas, 1999). The methodology follows the main steps: generate feasible chemical structures, estimate the thermo-physical properties through property models, and select the molecules that match the desired target properties (reject the molecules that do not match). Various algorithms have been proposed:

- Mathematical programming: all the steps of the generate-predict-select procedure are performed simultaneously. The molecular design problem is formulated as an optimization problem where the constraints are treated as mathematical equalities and/or inequalities and the performance indices are combined into an objective function, which is minimized through an appropriate numerical method. CAMD problems are usually formulated and solved as the Mixed Integer Non Linear Programming (MINLP) or Mixed Integer Linear Programming (MILP), when the constraint equations and the objective are linear equations. Examples of CPD as MINLP or MILP have been given by Duvedi and Achenie (1996), Churi and Achenie (1996), Camarda and Maranas (1999) and recently Zhang et al (2015). A difficulty for mathematical programming is the size and complexity of the mathematical programming models as well as the correct definition of constraint values to limit the search space. The advantage, however, is the only method where the optimal solution can be found;
- Generate and test approach (Gani, et al., 1991; Joback and Stephanopoulos, 1989; Pretel et al., 1994; Constantinou et al., 1996; Harper and Gani, 2000): this approach can handle the large size of the search space. All the steps of the generate-predict-select procedure are performed sequentially. For example, the complexity of the problem can be divided into the screening of pure compound properties and mixture properties. However, special combination rules are needed to avoid the combinatorial explosion, which occurs when the size of the problem becomes so large that computational time becomes too excessive.
- Decomposition methods (Solvason *et al.*, 2009; Chemmangattuvalappil *et al.*, 2010; Karunanithi *et al.*, 2005): the problem is decomposed in sub-problems and different tools are employed for each sub-problem;
- Stochastic optimization: this algorithm relies on the successive pseudo random generation of solution alternatives. Based on the solution after an iterative step the next solution is generated by introducing random permutations subject to probability functions. This algorithm does not require any gradient information allowing the easy specification of discontinuous properties as design goals. Marcoulaki and Kokossis (1998) and Ourique and Tell (1998) have proposed the

use of 'Simulated Annealing' method. Venkatasubramanian et al (1995) apply a 'Genetic algorithm' (Holland, 1975) based on Darwinian evolutionary theory;

- Special optimization techniques: Korb et al (2007) apply PLANTS (Protein-Ligand ANT System) which is based on ant colony optimization for structured-based drug design. Schneider et al (2009) design bioactive compounds through particle swarm optimization (PSO);
- Database search: database search may also be employed. It involves the selection of known compounds (Joback and Stephanopoulos, 1989; Modi *et al.*, 1996) from a database. The advantage of this approach is that it is easy to be applied when the search space is not very large, the best solution can be determined. However, this approach does not involve any generation of new molecules. Therefore, the databases are valid within a narrow product context.

The use of CAMD to identify novel molecules have been applied to design polymers (Satyanarayana et al., 2009), solvents for separation (Hostrup et al., 1999; Chemmangattuvalappil et al., 2010), refrigerants (Churi and Achenie, 1996; Cignitti et al. 2015), drugs (Koga et al., 1980); solvents substitution (Gani et al., 1991; Karunanithi et al., 2005) and many more.

Computer-aided mixture/blend design (CAM^bD)

Mixture/blend design problems can be defined as follows: given a set of chemicals and a set of property constraints, determine the optimal mixture and/or blend (Gani, 2004b). The chemicals to be mixed together are unknown, and also their relative compositions in the blend are not known. But the molecular structures of the candidate chemicals are known. Mixture design is similar to molecular design in the sense that both design problems combine building blocks in order to reach some *a priori* defined targets: in molecular design the building blocks are the groups (CH3, CH, OH,…) or atoms (C, H, O,…), while in mixture design, the building blocks are molecules. Mixture design involves the following:

- Mixture design requires the calculation of the relative amounts of chemicals to blend together (concentration) such that the target property constraints are matched;
- Mixture design implies the need to handle phase behavior issues, that is, miscibility/solubility between the ingredients.

According to Gani (2004) mixture/blend design is still a quite immature area, and there is just limited knowledge and know-how about a systematic approach for the design and verification of this type of chemical products. The main efforts have been directed to the design of solvent mixtures (Sinha *et al.*, 2003; Karunanithi *et al.* 2005) and recently, to blends (Yunus et al., 2014).

Formulated product design

Conte et al. (2011) developed a systematic methodology for design of homogenous formulated products, which has been adapted and extended by Mattei et al. (2014) for the design of emulsified products. These methodologies are based on CAMD and CAM^{b}D. The formulation design is decomposed into sub-problems (such as design of active ingredients, solvent mixture and additives), each sub-problem is solved separately using appropriate techniques. For instance, the design of active ingredients can be performed through CAMD or database search, while the design of solvent mixtures can be performed through CAM^{b}D.

Device and functional products design

Fung and Ng (2003) developed databases for synthesizing and developing a manufacturing process for pharmaceutical tablets and capsule that help for material selection, mechanistic models for particle strength, heuristics for equipment selection and equipment models for solids processing equipment. Morales-Rodriguez and Gani (2009) proposed a knowledge base containing data related to fuel-cells and microcapsule for control release design and multi-scale modeling approach to design products and to study behavior of products. These methodologies are based on the "define target- match target" paradigm. They employ the reverse design technique. That is, the product needs (defined by consumers or companies) are known and they are converted into a set of target properties, this set of properties are the constraints to be used to determine a set of promising candidates generated based on a specific algorithm of each product design methodology. Seider et al. (2015) proposed design procedures that combine model-based and experimental-based approaches for designing device and functional products.

2.5 Issues and needs

The issues and needs related to computer-aided product design problems are many and diverse. They are organized below under the following generic titles.

2.5.1 Product representation

Each type of CP has different way to represent the product structure and compound formula. In a single molecule product, the structure of the molecule is composed of different fragments (such as n-decane consists of 2 groups of CH_3 and 8 groups of CH_2). A blended product composes of a combination of more than one single molecule product mixed together, such as a liquid mixture of n-decane and n-dodecane. This way, the chemical system representing the product is simplified, thereby leading to a good understanding of the product and appropriate property models that can be reliably used. For example, in order to design a single molecular product, the property models should be able to estimate a wide range of compounds based on their molecular structures. Section 3.1 in chapter 3 gives more details about the representation of CPs.

2.5.2 Product candidate generation

The representation of CPs defines how candidates are generated. For example, the fragments of CH3, CH2, and CH can be combined in several ways to form feasible molecules based on the rules to control the structural feasibility of the generated molecules. Section 3.1 in chapter 3 gives more details about the generation of product candidates.

2.5.3 Problem definition

The reliability of a solution to a product design problem depends on the problem definition. This step consists of identifying the needs for a specific product, and relating these needs to physicochemical properties. There is the necessity of developing knowledge base systems that may guide the chemical product designer to convert the problem representation space from customer needs to technical specifications, as well as to specify their target values for a large range of chemical product design (Harper, 2000; Gani, 2004; Conte et al., 2011; Mattei et al., 2014). Costa et al. (2006) claim this is to be relevant for improving the understanding of the relationship between product performance, product composition, ingredient properties, processing variables and usage variables. In this work, product information that could be useful in the design and verification of a very wide range of chemical-based products, are collected and stored in the knowledge base, for easy retrieval and use when necessary.

2.5.4 Modeling

Modeling is a key step in the solution of all computer-aided product design problems.

- Properties and Property models: Property models are the key in the solution of all computer-aided product design problems since these property models are used to estimate properties of product candidates and eventually help to identify the optimum candidates which are heavily relied on the accuracy of the models. Property models are usually developed from regression analysis over a set of experimental data of compounds. The development of property models involves theory/hypothesis definition, model equation solving, validation of model against experimental data, and modification of theory/model parameters if required. From the cyclic process of property prediction models, it can be said that the accuracy of a model is affected by the uncertainties, which can arise from deficiency in theories or models and their parameters, and insufficient of knowledge of the systems (Kontogeorgis and Gani, 2004).The most significant limitations to the use of property models are associated with the unavailability of model parameters and the accuracy of prediction. If model parameters are not available for a product candidate, this molecule has to be discarded, since its properties cannot be estimated. The major need in this area is to extend the application range of existing property models, improving their performances, and, if necessary, develop new reliable property models. In this work, available models for the estimation of target pure component and mixture thermo-physical

properties are adopted and implemented into a model library in order to enhance their application ranges through the use of computer-aided tools.

- Multiscale modeling: It is necessary to organize time/length scales and complexity levels in some product and process engineering problems: first, understand and describe the phenomena and the properties at nano-, micro- and meso-scales; second, understand the relationships between the different scales
- Multidisciplinary modeling: The systematic frameworks for product design should take into account not only product structure and functionality, but also the manufacturing, management, sales and marketing, finance and economics (Gani and Ng, 2015). Multidisciplinary approaches need to be developed in response to the increasing environmental, societal and economic requirements and to the transition towards sustainability, that is, environmental protection, security, societal demands, and business including better conversion and selectivity of raw materials and energy for consumer desired product quality (Charpentier and McKenna 2004).

2.5.5 Methodologies

One of the main research challenges in the context of chemical product design is the development of systematic procedures, with related workflow and dataflow, where computer-aided tools are employed for a first screening of thousands of candidate, saving the valuable experimental resources for focused experiments. Although many computer-aided product design methodologies have been proposed, most of them can be applied only for the specific problem type for which it was developed. The key that allows wider application ranges for solving each type of product design problem is to store the corresponding product design work-flows, data-flows, tools, models and calculation algorithms in the template library in a format that is easy to maintain and update. The template is able to generate different versions of the parent template by modifying models, adding more product information, adding new models and many more. The template can be used when the design product has the same nature with the parent template (such as blend design template is created for gasoline blend problems but it can be used for diesel blend problems). It is very useful when there is a lack of available design methodologies for unknown or new products.

2.5.6 Computer-aided tools

As more complex the product design problem systems are, as more time and efforts are needed for their solution. Therefore, the use of software tools is necessary. The tools should be systematic but flexible, simple but accurate. A combination of tools is needed for design of some products, such as, property prediction tools is used to calculate product properties, optimizer is used to formulate and solve the product optimization problem. Therefore, in this PhD project, the computer-aided tools for CPD are created with the options that are able to work with other product design tools.

2.5.7 Systematic frameworks

The solutions of chemical product design problems require different methodologies and tools such as, databases, property models, design algorithms, computer-aided tools

many more steps. The development of such methods and tools is almost as important as the integration of them into a systematic framework that provide the work-flows and dataflow of the methodologies through the use of these computer-aided tools, at the same time, allows inter-changes of information, data and results. The structure of such a framework should be flexible, so that new models can be added, and new classes of products can be designed by using a common workflow. In addition, a user-friendly interface is required, so that the software can be employed for industrial application, as well as for effective teaching of chemical product engineering.

3 COMPUTER-AIDED MODEL-BASED FRAMEWORK FOR CHEMICAL PRODUCT DESIGN AND ANALYSIS

Design methodologies and tools for CPD play important roles in their integration to efficiently solve a wide range of product design problems, involving a wide range of issues, and therefore, needing different types of calculations. In this chapter, the models, data, algorithms and tools that are needed to solve a wide range of CPD problems are described as parts of a systematic framework.

The objective of the systematic framework (adapted from Seider et al. (1999)) is to efficiently manage the information related to CPD as illustrated in **Figure 3.1** in terms of the options needed for product simulation. A product designer defines the CPD problem by specifying the product type and product needs or product function; the knowledge base that stores product information and assists in the translation of needs into physicochemical properties (target properties) that require relevant property models and their parameters. In case the needed property models are not available in the model library, the product designer needs to import their models or modify existing models to match his/or her design objectives and implement them into a model library. When models are available, a product design template is used to guide the product designer to systematically design the product by following the needed actions. Each step of the design methodologies requires different work-flow/data-flow, models, algorithms and computer-aided tools. For example, in order to solve a MINLP problem, one would need a solver that should be available in the solver library. When promising product candidates that satisfy the predefined target properties are identified, they are verified by experiments and/or more rigorous models to ensure accuracy of property models as well as to verify some target properties (such as scent or appearance) which are not able to be modeled.

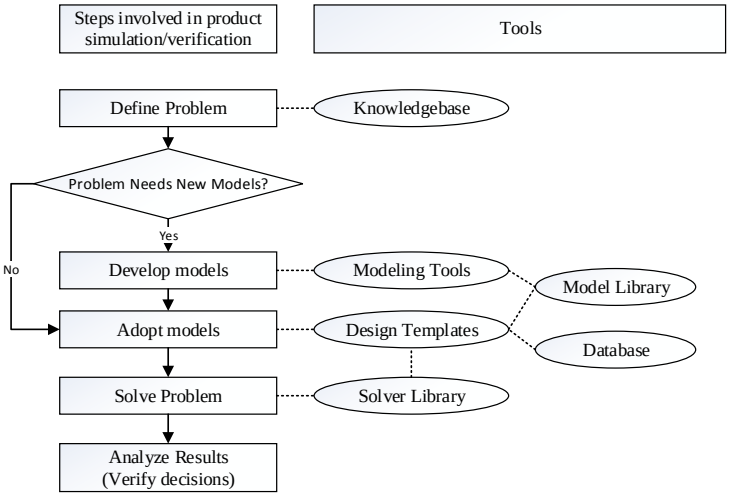


Figure 3.1. The steps involved in product simulation

3.1 Product structure representation and generation

In order to design each class of CP mentioned in Chapter 2, it is necessary to represent the products as well as generate product candidates with respect to how a product is represented.

3.1.1 Single molecule products

3.1.1.1 Molecular structure representation

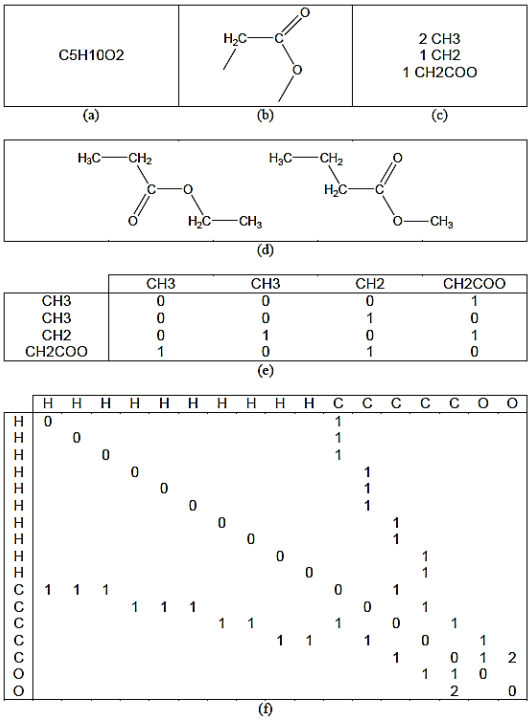
All CAMD techniques need to employ some form of compound representation in order to use molecular structure information for property estimation. In general, the estimation methods used for predicting properties of the designed molecule(s) decide the level of detail needed for the molecular structural information and the chemical representation method to use. Several different representations of molecules are highlighted in **Figure 3.2**:

- Atomic representation: this approach is the simplest form of a molecule that expresses information about the proportions of atoms that constitute a particular molecule such as $C_5H_{10}O_2$ can be used to represent Ethyl propionate (see **Figure 3.2a**). This approach is used when the property models need only information about the content of each type of atom in the molecule. For example, some correlations for estimation of higher heating value only need the weight percentages of elements(C, H, O, N, and S). This approach does not provide

information about the bonds in the molecule, however, it is possible to calculate the bond configurations if the valence of each elements are provided;

- Fragment: this approach gives information about how the atoms are connected, how many free connections the groups of atoms have and where (on which atom) they are located, as shown in **Figure 3.2b**. A molecular fragment can be grouped and defined by the number and types of atoms in each fragment. It gives information about the connectivity of the structure of the molecule (see **Figure 3.2c**) as well as represents isomers as shown in **Figure 3.2d**.

In order to provide information about structural groups and their connectivity to a computational tool, various methods exist (Raman and Maranas, 1988; Churi and Achenie, 1996; Maranas and Floudas, 1994). One of the most versatile and manageable methods is the adjacency matrix. An adjacency matrix is a square symmetrical matrix with rows and columns representing the atoms (or groups of atoms) in the molecule and containing zeroes and non-zeros indicating bonds and the absence of bonds. The adjacency matrix can be at fragment level (see **Figure 3.2e**) or at atomic level (see **Figure 3.2f**).



	CH3	CH3	CH2	CH2COO
CH3	0	0	0	1
CH3	0	0	1	0
CH2	0	1	0	1
CH2COO	1	0	1	0

(e)

	H	H	H	H	H	H	H	H	H	H	C	C	C	C	C	O	O
H	0										1						
H		0									1						
H			0								1						
H				0								1					
H					0							1					
H						0						1					
H							0						1				
H								0					1				
H									0					1			
H										0				1			
C	1	1	1								0		1				
C				1	1	1						0		1			
C							1	1			1		0		1		
C									1	1		1		0		1	
C													1		0	1	2
O														1	1	0	
O															2		0

(f)

Figure 3.2. Forms of molecular structure representation (Harper et al., 2000)

The size of the search space for generation of molecule candidates depends on the number of fragments, which are building blocks that could be connected and the fragment vectors that represent complete molecular structures.

3.1.1.2 Generation of molecule candidates

Molecule candidates can be generated in several ways (Gani et al., 1991; Joback and Stephanopoulos, 1989; Duvedi and Achenie, 1996; Constantinou et al., 1996). The number of fragments (building blocks) can be divided into 2 classes:

- First-order group: these fragment groups are based on UNIFAC groups (Constantinou and Gani, 1994). For example, methylbutane contains three first order groups (3 CH2, 1 CH2 and 1 CH);
- Second-order group: these groups are created regarding the effect of isomers (Marerro and Gani, 2001). For example, methylbutane contains one second order group (1 (CH3)2CH) in order to ensure that there are 2 CH3 branches in the end of methylbutane molecule.

For the hybrid generate & test based CAMD algorithm, the molecule candidate generation can be formulated as: Given a basis set of functional groups and the value of minimum and maximum numbers of groups allowed (specified) in the molecule and generate all the molecular structures that satisfy only structural constraints. The generated molecules are screened with respect to property constraints that can be calculated using first order and second order groups (Constantinou and Gani, 1994; Marrero and Gani, 2001). The molecule that satisfies specified property constraints is added to a collection of promising molecule candidates.

For the mathematical programming algorithms, the molecule candidate generation can be formulated as: Given an objective function, a basis set of functional groups and the value of minimum and maximum numbers of groups allowed (specified) in the molecule and generate all the molecular structures that satisfy structural constraints, property constraints and/or process constraints.

Structural constraints (adopted from Churi and Achenie (1996))

G_1 = {i | i is a first-order group};
G_2 = {j | j is a second-order group};
ID = {id | id is the ID number of each groups}.
Several binary variable representations are adopted in this model. Binary variable y_{i_1,id_1,i_2,id_2} denotes whether group i_1 with id id_1 (i_1, id_1) is connected to group i_2 with id id_2 (i_2, id_2), where i_1, $i_2 \in G_1$; id_1, $id_2 \in ID$. In this formulation, different bond type are considered within the structure of first-order groups, and all the second-order groups in Marrero and Gani (2001) are the connection of first-order groups using single bonds.

$$y_{i_1,id_1,i_2,id_2} = \begin{cases} 1 & \text{group } (i_1,id_1) \text{ is connected to group } (i_2,id_2) \\ 0 & \text{otherwise} \end{cases}$$

Binary variable z_{i_1,id_1} is used to describe the existence of group (i_1, id_1).

$$z_{i_1,id_1} = \begin{cases} 1 & \text{group } (i_1, id_1) \text{ exists in the molecule} \\ 0 & \text{otherwise} \end{cases}$$

Through classification of the different structural groups on the basis of their valence (number of free attachments), the octet rule provides a simple relation for the structural feasibility of a collection of groups (Odele and Macchietto, 1993).

$$\sum_{i \in G_1} \left(2 - v_i\right) n_i^{(1)} = 2q \tag{3.1}$$

In **Eqs. (3.1)**, $n_i^{(1)}$ is the number of first-order group i in the target molecule, v_i is the valency of group i, q is assigned the value of 1, 0 or -1 for acyclic, monocyclic or bicyclic groups, respectively.
In Churi and Achenie (1996), **Eqs. (3.2)** – **(3.6)** are added to ensure that only one molecule is formed.

$$\sum_{i_2 < i_1} \sum_{j_2} y_{j_1,i_1,j_2,i_2} + \sum_{j_2 < j_1} y_{j_1,i_1,j_2,i_1} \geq w_{j_1,i_1} \quad \forall i_1 > 1, j_1 > 1 \tag{3.2}$$

$$\sum_{i \in G_1} n_i^{(1)} + \sum_{j \in ID} \sum_{i \in G_1} w_{j,i} = n^{\max} \text{ and } w_{1,1} = 0 \tag{3.3}$$

$$w_{j_1,i_1} \geq w_{j_2,i_2} \quad \forall i_1 > i_2; i_1, i_2 \in G_1; j_1, j_2 \in ID \tag{3.4}$$

$$w_{j_1,i_1} \geq w_{j_2,i_1} \quad \forall j_1 > j_2; i_1 \in G_1; j_1, j_2 \in ID \tag{3.5}$$

Additional constraints may be placed on the number ($n_i^{(1)}$) of groups i to keep it within lower and upper bounds, n_i^L and n_i^U, respectively.

$$n_i^L \leq n_i^{(1)} \leq n_i^U \quad \forall i \in G_1 \tag{3.6}$$

Another constraints may be imposed on the total number of groups making up a molecule.

$$n^{\min} \leq \sum_{i \in G_1} n_i^{(1)} \leq n^{\max} \tag{3.7}$$

The adjacency matrix of target molecule can be established as shown in **Table 3.1**.

In the adjacency matrix, same groups with the same ID (diagonal) cannot be connected.

$$y_{i_1,id_1,i_1,id_1} = 0 \quad \forall i_1 \in G_1, id_1 \in ID \tag{3.8}$$

If group (i_1, id_1) connects to group (i_2, id_2), then (i_2, id_2) must connect to (i_1, id_1).

$$y_{i_1.id_1,i_2,id_2} = y_{i_2,id_2,i_1,id_1} \quad \forall i_1, i_2 \in G_1; id_1, id_2 \in ID \tag{3.9}$$

The constraints between binary variables z and y are shown in **Eqs. (3.10)** and **Eqs. (3.11)**.

$$\sum_{i_2 \in G_1} \sum_{id_2 \in ID} y_{i_1,id_1,i_2,id_2} = v_{i_1} z_{i_1,id_1} \quad \forall i_1 \in G_1, id_1 \in ID \tag{3.10}$$

$$\sum_{id_1 \in ID} z_{i_1,id_1} = n_{i_1}^{(1)} \quad \forall i_1 \in G_1 \tag{3.11}$$

Table 3.1. Adjacency matrix of target molecules

Groups		i_1	i_1	…	i_1	i_2	i_2	…	i_2	…	i_k	i_k	…	i_k
	ID	1	2	…	$n_{i_1}^U$	1	2	…	$n_{i_2}^U$	…	1	2	…	$n_{i_k}^U$
i_1	1	0		…				…		…			…	
i_1	2		0	…				…		…			…	
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
i_1	$n_{i_1}^U$			…	0		$y_{i_1,n_{i_1}^U,i_2,2}$	…		…			…	
i_2	1			…		0		…		…			…	
i_2	2			…	$y_{i_2,2,i_1,n_{i_1}^U}$		0	…		…			…	
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
i_2	$n_{i_2}^U$			…				…	0	…			…	
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
i_k	1			…				…		…	0		…	
i_k	2			…				…		…		0	…	
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
i_k	$n_{i_k}^U$			…				…		…			…	0

The other equations in the structural constraints restrict the number of second-order groups ($n_j^{(2)}$) from the adjacency matrix. For any second-order group *J*, constraints can be established based on its chemical structure (connection of first-order groups) to obtain $n_J^{(2)}$ as **Eqs. (3.11)** and **(3.12)** shows. $N_B{}^J$ is the number of bonds in second-order group *J*, b^J is binary variable, *T* is an integer parameter, and it depends on the structure of the second-order group (examples are listed below). *M* is a big number for big-M method. In Big-M method, appropriate value of *M* should be selected. The value of *M* should be the smallest values that work in the context of the model, because large values of *M* can cause branch-and-bound solvers to make slow progress solving the MIP model. In this formulation, the value of *M* = 20, because in all second-order groups, the number of bonds never larger than 20.

$$N_B^J - M\left(1 - b_{id_1,id_2,\ldots,id_{N_B}}^J\right) \leq \sum_{\substack{\text{if group } (i_1,j_1) \text{ and } (i_2,j_2) \text{ are} \\ \text{connected in 2nd group } J}} y_{i_1,j_1,i_2,j_2} \leq N_B^J - 1 + M\left(b_{id_1,id_2,\ldots,id_{N_B}}^J\right) \quad \forall J \tag{3.12}$$

$$n_J^{(2)} = \frac{1}{T} \sum_{(id_1,id_2,id_{N_B})} b_{id_1,id_2,\ldots,id_{N_B}}^J \quad \forall J \tag{3.13}$$

Equations of several second-order groups are listed below as examples.

- $(CH_3)_2CH$:

$$2 - M\left(1 - b_{id_1,id_2,id_3}^{(CH_3)_2CH}\right) \leq y_{CH,id_1,CH_3,id_2} + y_{CH,id_1,CH_3,id_3} \leq 1 + M b_{id_1,id_2,id_3}^{(CH_3)_2CH} \quad \forall id_1, id_2, id_3 \in ID \tag{3.14}$$

$$n^{(2)}_{(CH_3)_2CH} = \frac{1}{2}\sum_{id_1\in ID}\sum_{\substack{id_2\neq id_1\\ id_2\in ID}}\sum_{\substack{id_3\neq id_2\\ id_3\in ID}} b^{(CH_3)_2CH}_{id_1,id_2,id_3} \quad (3.15)$$

In **Eqs. (3.14)**, if and only if $y_{CH,id_1,CH_3,id_2} = y_{CH,id_1,CH_3,id_3} = 1$, there exist a second-order group $(CH_3)_2CH$, and $b^{(CH_3)_2CH}_{id_1,id_2,id_3} = 1$. Since the two CH_3 groups in group $(CH_3)_2CH$ are counted twice as they have different ID, the number of the second-order group $(CH_3)_2CH$ equals to ½ times the summary of $b^{(CH_3)_2CH}_{id_1,id_2,id_3}$ as **Eqs. (3.15)** shows.

All second-order groups are formulated in this way to obtain their number from the adjacency matrix. These constraints do not need to be modified for different problems. Thus, these second-order constraints can be stored separately for all CAMD problems.

Property constraints

The property constraints are represented in **Eqs. (3.16)**. P is the set of all target properties of the molecule. All the target properties should be in its given range [p_k^L, p_k^U].

$$p_k^L \le p_k \le p_k^U \quad \forall k \in P \quad (3.16)$$

The target properties p_k may be obtained from the molecular structural variables or the combination of other properties. Constantinou et al. (1996) proposed a classification of properties as primary (pure component properties that can be determined only from the molecular structural variables as **Eqs. (3.17)** shows), secondary (pure component properties that are dependent on primary properties) and functional (pure component properties dependent on temperature and/or pressure).

$$p_k = \sum_{i\in G_1} n_i^{(1)} p_k^{(1)} + \sum_{j\in G_2} n_j^{(2)} p_k^{(2)} \quad \forall k \in P \quad (3.17)$$

Process model and other constraints

The process model and other constraints contain continuous and discrete variables. These constraints integrate the product design problem with process design problem, economic model, etc.

3.1.2 Blended products

A blended product is composed of compounds (more than one) that are mixed together and form homogeneous phase. The structure of these products can be liquid products (such as fuel blends) or solid products (polymer blends). Different compounds have different properties that contribute to the bulk properties. In the case that one compound (a single molecule product) cannot satisfy all product functions (or target properties), the single molecule is then blended with other compounds. Therefore, blended products are presented as:

- Main ingredient (MI): MI is a compound that is mostly found in a blend. In some cases such as fuel blends, the MI can be mixtures of hydrocarbon compounds which are called MIs:
- Additive candidates: additives are blended with MI(s) in order to improve the properties of MI able to satisfy all target properties.

A list of additive candidates can be generated by specifying target properties of feasible additives and using CAMD methods to generate and screen possible candidates that satisfy all target properties of feasible additives. For example, the additive candidates for lubricant blends can be generated by specifying the number of groups of esters, parafins, iso-paraffins and naphthenes that should be presented. By employing the hybrid generate & test based CAMD algorithm, thousands of lubricant additives are generated and their properties are calculated. The lubricant additives that satisfy all specified properties are collected as a set of additives to be blended with MI.

3.1.3 Formulations

They constitute a class of 'consumer oriented chemical products'. They are formed by several ingredients (from 5 to 20). They can provide for several functions, and can have different forms (powder, solution, emulsion). For instance, a sunscreen lotion has the function of blocking the UV radiation, avoiding skin cancer, slowing the skin aging. Sunscreens can have the form of creams (emulsions) or solutions of oils, which can also be sprayed through a nozzle. The structure of the product defines how product constitutions are combined and the product constitutions depend on product functions. The constitutions of the liquid-base formulated product can be classified into three types:

- Active ingredient (AI), or key ingredient: the AI is the ingredient that provides the main function of a formulation (also referred to as 'activity'). Since a formulation can provide several functions, more than one AI may be present in a single formulated product. AIs of homogenous formulated products, for example, an insect repellent lotion. It is consisted of compounds that have the functions of protecting the skin from sunburns and skin cancer, but also of preventing the skin aging. In the case of emulsified formulated products (such as a detergent), AIs are emulsifier that make the products form emulsion;
- Solvents: the solvent of the product can be pure component or mixtures of compounds. They are used for product compatibility purposes such as delivery of AI to the target area (such as deliver an AI of insect repellent lotion to theskin). The class of compounds that are chosen to be a solvent is based on an AI. For example, water soluble solvents will be used if the AI is well known to be water soluble. The highest concentration of compounds presented in the product are solvents;

- Additives: additives can be used to enhance product qualities such as the solubilization of AIs, the product's stability, product spread ability on surfaces, avoiding a microbial growth, sensorial factors. The concentration of additives presented in the product is very low compared to AI and solvents.

3.1.4 Devices

For chemical devices or functional products, they include the type and amount of the key ingredients, and the structure with which these ingredients are configured:

- Key ingredients: the key ingredients refer to the components that are essential for achieving the desired outcome (the product performance) (Seider et al., 2009). For example, a ascorbic acid (Vitamin C) tablet. The product performance is to deliver the vitamin C to be absorbed into the body quickly and should not break up on routine handling as well as release the designed amount of vitamin C to the body. Therefore, the key ingredients are: the vitamin C and materials to be made as a tablet to deliver the vitamin C to the body (Fung et al., 2003);
- Product structures: the nature of the selected key ingredient and the product performance define the product structure, size and its configuration. For example, the vitamin C is fastly dissoluted in water and the standard disintegration time of the vitamin C tablet is 30 min. The vitamin C is water soluble. Therefore, the size of the tablet should be round 5 μm. The tablet is used at the room condition, therefore, the nonhygroscopic tablet is select to cover the vitamin C because it is stable at the room condition. The process to select the size, product structure and its configuration can be done through the use of models to predict the associated phenomena such as mass transfer and chemical reactions in order to fine the optimal specification to reach the product performance.

3.2 Problem definition

Problem definition is the first step in order to design all chemical-based products. This step identifies: what type of products should be designed with respect to product needs, what are target properties should be considered, what kind of compounds should be in the product formula with respect to target properties, and what are experiments that should be performed to verify the product formula. This step requires knowledge of experts and review of published articles, which consumes a lot of time and human resources. Therefore, there is a need for the development of a reliable knowledge base with information-data from publications, patents, patented products, insight and common sense. The product designers should be able to use the knowledge base and add his/her own specific ideas to extend the knowledge base.

3.2.1 Knowledge base

A reliable knowledge base should offer the following options:

- The identification of the product needs/ functions;
- The identification of a set of physicochemical properties that are used to define the set of performance criteria;
- The setting of the constraint values on the target properties. These values may come from patented products which are taken as references for the setting of the constraint values or the product designer can also decide to improve the constraint values, in order to improve the existing product. The constraint values define the range or numbers of product candidates. The sensitivity analysis on constraint values are needed to be done in order to identify the most sensitive target constraints that define the product design results. For example, blending a conventional gasoline (main ingredient) with additives. The target properties, higher heating value (HHV) ($\geq$ 40 kJ/mol) and liquid density at 20 ^{0}C lies between 0.750 g/cm^3 and 0.775 g/cm^3. The HHV of the main ingredient is 45 kJ/mol and the density is 0.750 g/cm^3. In this case, the density limits the number of additive candidates because the narrow constraint range;
- All the other type of information, which could be useful in the design and verification of CPs such as an identification of the product qualities to enhance with the addition of additives for formulated products.

Product information is collected for supporting the decision and choices required during the design and analysis. The information is managed through ontology of product information system with respect to: product types, product needs, relevant chemicals, translation of product needs into target properties and suggested experimental verification as shown in **Figure 3.3**. For example, under "blends" a sub type, "gasoline blend" is stored. Each product subtype has different consumer needs, that is, the gasoline blend has to have the ability to be burned. Each needs are involved with: several classes of chemicals (such as alkane and alcohols are combustible chemicals); several target properties and target values which are divided into four main types (primary property, secondary property, functional property and mixture property); and suggested experimental tests that should be performed in order to verify the product needs.

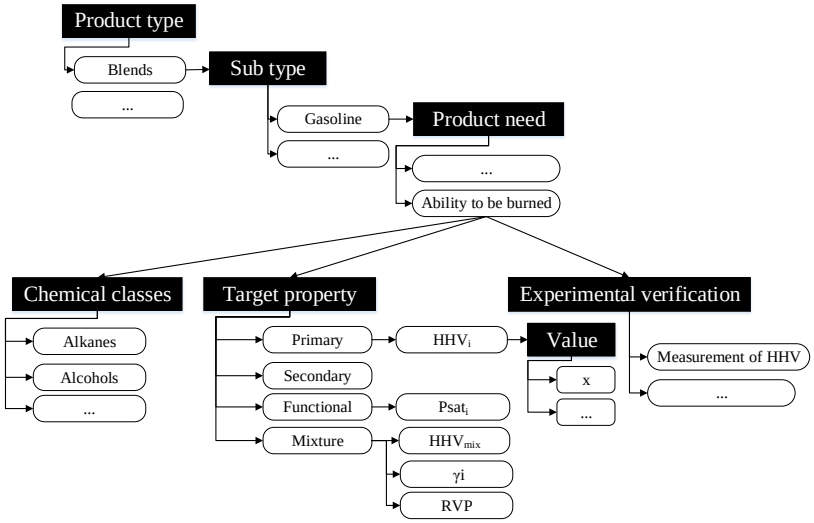


Figure 3.3. VPPD-Lab knowledge base structure

3.2.1.1 Product needs and translation of needs into target properties

After the product type is identified, the product needs are identified. The identification of product needs and translation of the needs into target properties of each product type are described below.

Single molecule products

Single molecule products are various types of chemicals (such as solvents, refrigerants and active ingredients). **Table 3.2** gives examples of the essential needs for solvents based on their process application.

Table 3.2. List of solvent-based product and their needs

Solvent product	Need	Target property	Reference
Solvent for liquid-liquid extraction	Liquid state at operating temperature	boiling point, melting point	Harper, P. M., 2000
	Miscible with the solute that is to be separated from the feed solution and immiscible with the feed solution	density, solubility parameter	
	High efficiency	separation factor, solvent capacity, selectivity, solvent loss	
	Safety and environmentally friendly	Open cup flash point, global warming potential, human toxicity, flash point, ozone	
	Easy to be recovered	boiling point, vapor pressure, no azeotrope (choice of solvent), miscibility	
Solvent for solid liquid extraction	Liquid state at operating temperature	boiling point, melting point	Harper, P. M., 2000
	Miscible with the solute	solubility parameter	
	Safety and environmentally friendly	Octanol/water partition coefficient, open cup flash point, global warming potential, human toxicity, flash point, ozone	
Solvent for a pharmaceutical compound	Liquid state at operating temperature	boiling point, melting point	Harper, P. M., 2000
	Miscible with the solute	solubility parameter	
	Safety and environmentally friendly	Octanol/water partition coefficient, human toxicity	
Single solvent for extractive distillation	Selectivity dissolve solute	Selectivity	Harper, P. M., 2000
	No phase split	Solubility , miscibility	
	Safety and environmentally friendly	Global warming potential, human toxicity, flash point, ozone depletion potential, vapor pressure	
	No azeotrope	Choice of solvent	

In solvent design for liquid-liquid extraction, a solvent should not evaporate or form a solid particle during the extraction, therefore, normal boiling point should be higher than the operating temperature, as well as, the melting point should be lower than operating temperature. The density must be different from the feed solvent (the solvent that the solute is dissolved) (the ratio of the density values at the operational temperature must be at least 1.05) (Harper et al., 2000). Solvent loss, separation factor, solvent capacity, and selectivity must be high as possible. Regarding EH&S properties, open cup flash point of the solvent should be above the operating temperature. When the solvent dissolves the solutes and splits phase from the feed solvent, it should be easy to be recovered the solvent by the simple distillation as well. For solvent design of solid-liquid extraction and solvent for a pharmaceutical compound, EH&S properties

are very important, therefore, octanol/water partition coefficient, human toxicity are used as indicators. For solvent design of extractive distillation, a solvent should not form azeotrope with the other compound in the mixture that needs to separate.

Blends

Blends are classified into sub-types such as gasoline blends, jet-fuel blends and diesel blends. **Table 3.3** list the product needs and target properties of fuel blends (gasoline blends, jet-fuels blends, diesel blends and lubricant blends).

Table 3.3. Product needs and their target property constraints of fuel blends

Need	Target Property	Gasoline	Jet-fuels	Diesel
Ability to be burned	Reid vapor pressure	*	*	*
Safety	Flash-point	*	*	*
Engine efficiency	Higher heating value	*	*	*
	Density	*	*	*
	Octane rating	*		
Consistency of fuel flow	Dynamic viscosity	*		
	Kinematic viscosity		*	*
	Melting point		*	
Stability	Miscibility	*	*	*
Environmental impacts	Oxygen content	*		
	Human toxicity (-logLC50)	*	*	*
	CO_2 emission		*	*

The gasoline blends (Yunus et al., 2014) must have: enough Reid vapor pressure to ensure that the fuel is sufficiently vaporized to be burned and the engine can start at the operating temperature; high higher heating value; very high flash-point; and low viscosity to continuously flow from the fuel tank to the combustion chamber. Furthermore, the gasoline blends must be stable, meaning that the blends do not evaporate easily; do not oxidize to form unwanted by-products, such as gums, sludge and deposits during storage; and must not split into two liquid phases.

For the jet-fuels blends (Gammon, 2004), the specifications of the higher heating value and environmental impacts are supposed to be high and low, respectively. However, the jet engines are operated at low temperature above the ground. Therefore, the jet-fuels blends should have: low Reid vapor pressure to prevent evaporative losses and fuel system vapor lock; high flash-point; high higher heating value to maximize the energy that can be stored in a fixed volume and provides the longest flight range; low melting point to prevent the wax crystals which is difficult to be pumped into turbine engines; and low greenhouse gas emissions.

The diesel blends (Garrett, 1994) must have: low Reid vapor pressure; high flash-point; low viscosity; and low environmental impacts as well as human toxicity.

Table 3.4 lists product needs and target properties of lubricant blends (Yunus et al., 2014). Basic performances of a lubricant are able to reduce friction between two moving surfaces and able to suspend and remove the impurities. Furthermore, engine oils must have the following criteria: low viscosity to flow at the surrounding temperature; low viscosity index (the extent of viscosity change with temperature) in order to be operated at a high temperature; high pour point to prevent the forming of microcrystals; and high flash-point but low volatility to prevent losses to the ambient by evaporation.

Table 3.4. Product needs and their target property constraints of lubricant blends

Need	Target Property
Ability to lubricate and prevent wear	Kinematic viscosity
Resist at high temperature	Viscosity index
Ability to flow at ambient Handling purpose	Pour-point Liquid density
Non-flammable	Flash-point
Low vaporization rate	Volatility

Formulations (homogeneous phase products)

Formulations (homogeneous phase products) (Conte et al., 2011) are divided into sub-types such as an insect repellent lotion and a UV sunscreen lotion (see **Table 3.5**). For an insect repellent lotion, AI must be able to prevent insect from a human skin. Solvent mixtures should: compose of water plus alcohol for safety and toxicology issues; have enough evaporation time (it should not be needed to apply the product often during exposure to mosquitoes); low toxicity; and low kinematic viscosity for good spray-ability. For a waterproof sunscreen lotion, AIs must protect sunburns, the risk of skin cancer, skin aging. The solvent mixtures should: have solubility parameters close to AIs; be able to dissolves AIs; and not have too high viscosity.

Table 3.5. Product needs and their target property constraints of formulations

Need	Target Property	Insect repellent lotion	Waterproof sunscreen lotion
Effectiveness to protect skins from insects	Choice of AI	*	
Material compatibility	Choice of solvent ,	*	*
Cosmetic properties (Odor)	Choice of additive	*	*
Durability	Evaporate time	*	*
Low toxicity	Human toxicity (-logLC50)	*	*
Stability	Hildebrand solubility parameter, miscibility	*	*
Good spray-ability	Kinematic viscosity Density	* *	* *
Protection of sunburns, skin cancer and skin aging	Choice of AI		*
Water resistance	Choice of AI and solvent		*

Formulations (emulsified products)

For emulsified products, the formulation depends on the composition (formulation) of compounds present in them and the temperature as outlined in the Kahlweit's fish phase diagram (Lin and Chen, 2004). This phase diagram can be used during the design and/or verification steps. However, experimental data are needed to generate this diagram. Therefore, Mattei et al (2014) applied the hydrophilic-lipophilic deviation (HLD) approach to identify the formulation of the products. Zero HLD indicates a hybrid area where an emulsion and a microemulsion may coexist and therefore an unstable system is expected. On the other hand, if a positive value is obtained, then a water-in-oil emulsion is favored, while a negative value indicates that an oil-in-water emulsion may be formed. The higher the absolute value of the HLD of the surfactant is, the more stable the emulsion formed is expected to be, since it is located further away from the unstable region identified by the hybrid domain. This method does not have the thermodynamic basis of the representation of the ternary phase diagram, but it can be used as a qualitative predictive model when the needed experimental data are not available. The target properties of a hand wash detergent and a tank cleaning detergent are listed in **Table 3.6**. In order to make the ingredients form an emulsion, the surface tension of the surfactants should be lower than water and oils which can be achieve by adding surfactants. The critical micelle concentration is defined as the concentration of surfactants above which micelles form and all additional surfactants added to the system go to micelles. Therefore, the critical micelle concentration should not be high in order to save the amount of surfactants added. pH of emulsified products should be around 4 to 7 in order to avoid skin damage. Viscosity should be low for easy applications and spread ability. The flash-point should be high for safety purposes.

Table 3.6. Product needs and their target property constraints of a hand-wash detergent and a tank cleaning detergent

Need	Target Property	Hand-wash detergent	Tank cleaning detergent
Foam-ability	Surface tension	*	*
	Critical micelle concentration		
Non irritability of skin	pH	*	
Wetting of the substrate	Surface tension	*	*
Dissolution of the dirt	Solubility parameter	*	*
Suspension of the dirt	Hydrophilic-lipophillic balance	*	*
Spread ability	Molar volume	*	
	Viscosity	*	*
	Coud point	*	
	Krafft temperature	*	
Stability	Hydrophilic-lipophilic deviation	*	
Good stability to temperature and dilution	Coud point		*
High safety	Flash point		*

Devices

The main need for devices is the ability to perform the desired performances such as delivery of a certain amount of drugs into the blood -at a certain period of time. The product structure defines the product formula. **Table 3.7** gives an example of the needs for some device products (Morales-Rodriguez and Gani, 2009). The values of the target properties depend on the application of the devices. For example, for the controlled release, % release of the drug should not be high for some drugs but % release should be constant.

Table 3.7. Product needs and their target property constraints of a micro-capsule controlled release

Product	Need	Target Property
Pesticide uptake	Ability to uptake pesticide from water droplet to leaf	Relative uptake of active ingredient
Microcapsule controlled release of active ingredients	Ability to delivery drugs into the blood with a certain amount at a certain period of time	% Release

3.3 Property modeling

The identified product properties play a fundamental role in the design and development of CPs. An integrated experiment-modeling approach is usually recommended, where in a first model-based stage, property models are used to estimate the desired set of properties and a set of promising candidates are identified. In a second experiment-based stage, the properties are verified and the formulated product is further improved, if necessary. Therefore, a database of collected experimental data, supported by reliable mathematical models for prediction of thermo-physical properties is of fundamental importance.

3.3.1 Property classification

The chemical properties are classified in terms of a hierarchical order as:

- Primary properties: these are single value properties of the pure compound. Every molecule is characterized by a single value of these properties such as normal boiling point, critical temperature, critical pressure and many more;
- Secondary properties: these properties of the pure compound can be calculated from the primary properties, for example, acentric factor is a function of critical temperature, critical pressure and boiling point;
- Functional properties: these properties depend on temperature or pressure of the system where the pure compound is presented. For example, vapor pressure is the function of temperature, critical temperature, critical pressure and acentric factor;
- Mixture properties: these properties depend on mixture composition as well as temperature and/or pressure (or a defined mixture state). They represent the bulk property of the mixture, for example, liquid density, liquid viscosity and liquid

thermal conductivity of the liquid mixture. These properties also include phase equilibrium-related properties such as the activity or fugacity coefficients of each compound presented in the mixture.

- Performance-related properties: these properties are related to the performances of the product, such as the evaporation rate of the solvent and the stability of the blends or emulsions. These require the above property models embedded into a process/product performance model, which usually involve multiscale modeling.

3.3.2 Database

Chemical databases are fundamental tools for the solution of chemical product design problems. In this PhD dissertation, chemical data for a very wide range of chemicals found in different databases are divided into sections representing each specific type of chemicals.

Table 3.8 lists classes of compound and property data (primary, secondary, functional, and mixture properties) of lipids compounds, environmental-related compounds, emulsions, azeotropic compounds, solubility and related properties of complex chemicals (DECHEMA). Lipids compounds can be classified as nonpolar organic compounds that are insoluble in water. Collecting the available experimental data from different sources for the identified lipid compounds and their corresponding properties has been initiated by Tovar et al (2013) and completed within this PhD project. Environmental-related compounds database (Hukkerikar et al., 2012b) stores environmental related properties that are useful for life cycle assessment (LCA) and product-process sustainability analysis (Kalakul et al., 2014a). Emulsions database contain property data for design of emulsified product (Mattei et al., 2014). Azeotropic compounds database contains azeotrope experimental data to aid the design of azeotropic separation process in an easy, fast, reliable and predictive way (Gani and Bek-Pedersen, 2004). The solubility and related properties of complex chemicals database is created by systematically collecting published solid-liquid equilibrium data for a range of compounds with molecular weights starting near 100 g/mol. The largest molecules have carbon numbers near 40. The aim of the effort is to create a comprehensive collection of solid-liquid equilibrium data that is useful to the pharmaceutical, agrochemical, specialty chemical, and life sciences industries and for studies involving products and processes from these industries. This data collection will hopefully motivate the development of new and better property prediction models for solubility of complex organic chemicals so that products from these industries can be brought to market faster and at reduced costs. This is very important in product-process development since time to market may decide the success or failure of a product or a process that is designed to produce it (Kalakul et al., 2014b).

Table 3.9 lists classes of compound and property data of a wide range of compounds: normal fluids, polar compounds, polymers, electrolytes, amino acids, solvents, combustible compounds, and formulations. Formulations database stores experimental data for the design of formulations type homogeneous liquid products (Conte et al., 2011). Combustible compounds database stores lists of additives for blend design of jet-fuels, diesels, gasolines as well as lubricants (Yunus et al., 2014). Solvent database

stores information about common solvents that can be used for CPs. For example, information of Eastman n-butyl propionate can be found in the database. It is manufactured by ExxonMobil's and can be found in: coatings; cleaners; printing inks; and process solvents.

Table 3.8. VPPD-Lab databases (1)

Database	Compound	Number of compound	Data	Reference
Lipids	Alkanes, naphthalenes, aromatics, cyclic and poly-cyclic compounds, sulfides & mercaptans, silanes, phosphines, water and olefins, ethers, esters, ketones, aldehydes, anhydrides, nitro compounds, sulfonic compounds, azides, nitrates, nitrites, phophoric compounds, alcohols, acids, amides, oxilmes, oxides, nitriles, isocyanates, isothiocyanates, isocyanides, oxides and phophoric acids	330	*Primary and secondary properties*: melting point, boling point, critical temperature, critical volume, standard Gibbs free energy of formation, standard enthalpy of formation, standard enthalpy of fusion, specific gravity at 60 F, liquid volume, dipole moment and acentric factor; *Functional properties*: vapor pressure, liquid enthalpy, liquid viscosity, liquid thermal conductivity, ideal enthalpy, surface tension, latent heat, liquid density, vapor viscosity, vapor thermal conductivity; *Mixture properties*: binary data for vapor-liquid equilibrim (VLE), liquid-liquid equilibrium (LLE), solid-liquid equilibrium (SLE)	Diaz-Tovar et al., 2011; Cunico et al., 2013
Enviromental-related compounds	Alkanes, naphthalenes, aromatics, cyclic and poly-cyclic compounds, sulfides & mercaptans, silanes, phosphines, water and olefins, ethers, esters, acids, phosphoric compounds and furans	26,155	*Primary and secondary properties*: Fathead minnow 96-hr LC50, daphnia magna 48-hr LC50, oral rat LD50, bio-concentration factor, photochemical oxidation potential, global warming potential, ozone depletion potential, acidification potential, emission to urban air (carcinogenic) , emission to urban air (non-carcinogenic), emission to rural air (carcinogenic), emission to rural air (non-carcinogenic), emission to fresh water (carcinogenic), emission to fresh water (non-carcinogenic), emission to sea water (carcinogenic), emission to sea water (non-carcinogenic), emission to natural soil (carcinogenic), emission to natural soil (non-carcinogenic), emission to agriculrural soil (carcinogenic) and emission to agriculrural soil (non-carcinogenic)	Young et al., 2000; Hukkerikar et al., 2012b
Emulsions	UV-A absorbers, UV-B absorbers, UV filters, antioxidants, preservatives, aromas, co-surfactants, builders database, buffering agents, bleaching agents, colorants, anti-microbial agents and emollients	472	*Primary and secondary properties*: flash point, Fathead minnow 96-hr LC50, cost, critical micelle concentration, colud point, hydrophilic-lipophilic balance, ; *Functional properties:* density, liquid viscosity and surface tension; *Mixture properties:* binary data for liquid-liquid equilibrium (LLE)	Mattei et al., 2014
DECHEMA	The types of components found in the collection are those that can be constructed from two or more of the atoms: C, H, O, N, F, Cl, Br, I, S and P. The solutes have molecular weights starting near 100 g/mol and have from a few and up to about 40 carbon atoms. Inorganic compounds and polymers are not included.	353 solutes, 178 solvents	*Primary and secondary properties*: normal melting point, enthalpy of fusion, Hildebrand solubility parameter, octanol/water partition coefficient; *Mixture properties*: binary data and ternary data for solid-liquid equilibrium (SLE)	Kalakul et al., 2014b

Table 3.9. VPPD-Lab databases (2)

Database	Compound	Number of compound	Data	Reference
Normal fluids	Alkanes, naphthalenes, aromatics, cyclic and poly-cyclic compounds, sulfides & mercaptans, silanes, phosphines, water and olefins	1,664	*Primary and secondary properties:* acentric factor, critical temperature, critical pressure, critical volume,critical compressibility, melting point, boiling point,triple-point temperature and pressure, boiling point at specified pressure, liquid volume at 298.15 K, ideal gas enthalpy at 298.15 K, ideal gas Gibbs energy at 298.15 K, ideal gas entropy at 298.15 K, density, solubility parameters, van der Waals surface area and volume, radius of gyration, dipole moment, octanol/water partition coefficient, refractive index, molecular refraction, enthalpy of fusion, enthalpy of combustion and flash point temperature, relative permittivity; *Functional properties:* solid density, liquid density, vapor pressure, heat of vaporization, solid heat capacity, liquid heat capacity, ideal gas heat capacity, second virial coefficient, liquid viscosity, vapor viscosity, liquid thermal conductivity, vapor thermal conductivity and surface tension; *Mixture properties:* binary data for vapor-liquid equilibrium (VLE), liquid-liquid equilibrium (LLE), solid-liquid equilibrium (SLE), infinite dilution activity coefficients, enthalpies of mixing, partial molar enthalpies of mixing at infinite dilution, excess Gibbs energies, Henry's law constants, and mutual solubilities; ternary data for VLE, LLE, SLE, VLLE, enthalpies of mixing, and binodal data	Nielsen et al., 2001
Polar non-associating compounds	Organic & inorganic compounds	3,078		
Polar associating compounds	Alcohols, acids, ketones, ethers, aldehydes, esters, amines, halogens, peroxides, nitriles, anhydrides, amides, oximes, nitro compounds, mercaptans & sulfides, sulfonic compounds, cyanotic compounds and phosphoric compounds	2,355		
Multifunctional group compounds	Concentrated compounds and ionic compounds	4,445		
Polymers	week compounds, strong compounds, oxides and organics	23		
Electrolytes	Adrenal corticosteroids, androgens & Anabolic agents, estrogens, progestogens & progestins and cholesterols	124		
Amino acids		104		
Solvent		1,350	Manufactors, recommended products and potential substitution	Nielsen et al., 2001
Formulations	Pigments, insect repellents, UV-A blockers, UV-B blockers, anti-oxidants, polymers, water insoluble solvents, water soluble solvents, water insoluble alcohols, water soluble alcohols, esters, hair spray solvents, water, aroma compounds, preservatives, wetting agents and moisturizing agents	614	*Primary and secondary properties:* density, cost, Fathead minnow 96-hr LC50, dielectric constant, meltin point, evaporation time, solvents, non-solvents and solubility parameter; *Functional properties:* dynamic viscosity, kinematic viscosity, liquid volume, surface tension, vapor pressure; *Mixture properties:* binary data for solubility, vapor-liquid equilibrium (VLE), liquid-liquid equilibrium (LLE)	Conte et al., 2011
Combustible compounds (for fuel blends)	Alkane, cycloalkane, alcohols, olefins, olefin-alcohols, etoxy-alcohols, ethers, aldehydes, ketone, esters, acids, amines, amines, furans, alkynes, aromatics and halogenated compounds	1,725	*Primary and secondary properties:* melting point, boling point, critical temperature, critical volume, standard Gibbs free energy of formation, standard enthalpy of formation, standard enthalpy of fusion, standard heat of combusion, Fathead minnow 96-hr LC50, flash point, critical pressure and solubility parameters; *Functional properties:* liquid density, liquid viscosity, surface tension, vapor pressure and liquid heat capacity; *Mixture properties:* binary data for vapor-liquid equilibrim (VLE), liquid-liquid equilibrium (LLE)	Yunus et al., 2014

3.3.3 Property models

A database of collected experimental data, supported by reliable mathematical models for prediction of thermo-physical properties is of fundamental importance. A collection of the available property models is then presented according to the property model types:

- Pure compound property models;
- Mixture property models;
- Product performance models.

In addition, the property model for complex esters developed in this work is presented.

3.3.3.1 Collection of property models

Pure compound property models

Pure compound property models – Pure compound property models can be divided into three classes: primary property models, secondary property models, functional property models.

- Primary property models are: critical properties (see **Table 3.10**);
- environmental-related properties (see **Table 3.11** and **Table 3.12**);
- transport properties (see **Table 3.11**);
- combustion properties (**Table 3.11 and Table 3.12**).

These primary properties are used to calculate

- secondary properties (see **Table 3.13**)
- functional properties (see **Table 3.14**).

Details about property model equations and parameters for pure compound property models are given in:

- **Table 3.15** (for primary property model);
- **Table 3.16** (for secondary property models);
- **Table 3.17** – **3.18** (for functional property models).

Table 3.10. Property model of pure compounds – primary properties (1)

Property	Model	Function	Eq.
Normal Melting Point, T_m	Hukkerikar et al., 2012a	$f(GC)$	3.18
Normal Boiling Point, T_b	Hukkerikar et al., 2012a	$f(GC)$	3.18
Critical Temperature, T_c	Hukkerikar et al., 2012a	$f(GC)$	3.18
Critical Pressure, P_c	Hukkerikar et al., 2012a	$f(GC)$	3.18
Critical Volume, V_c	Hukkerikar et al., 2012a	$f(GC)$	3.18
Standard Gibbs Free Energy of Formation at 298 K, ΔG_f^{298K}	Marrero and Gani, 2001;	$f(GC)$	3.18
Standard Enthalpy of Formation at 298 K, ΔH_f^{298K}	Marrero and Gani, 2001;	$f(GC)$	3.18
Pitzer's Acentric Factor, ω	Marrero and Gani, 2001;	$f(GC)$	3.18
Enthalpy of Vaporization at 298 K, ΔH_{vap}^{298K}	Marrero and Gani, 2001;	$f(GC)$	3.18
Enthalpy of Vaporization at Tb, ΔH_{vap}^{Tb}	Marrero and Gani, 2001;	$f(GC)$	3.18

Table 3.11. Property model of pure compounds – primary properties (2)

Property	Model	Function	Eq.
Enthalpy of Fusion, ΔH_f	Marrero and Gani, 2001;	$f(GC)$	3.18
Liquid Molar Volume at 298 K, V_m^{298K}	Marrero and Gani, 2001; Rackett Modified Correlation;	$f(GC)$ $f(Tc, Pc, \omega)$	3.18 3.19
Liquid Surface Tension at 298 K, Sur^{298K}	Marrero and Gani, 2001;	$f(GC)$	3.14
Hansen Dispersive Solubility Parameter, δ_D^{298K}	Marrero and Gani, 2001;	$f(GC)$	3.18
Hansen Polar Solubility Parameter, δ_P^{298K}	Marrero and Gani, 2001;	$f(GC)$	3.18
Hansen Hydrogen-Bond Solubility Parameter, δ_H^{298K}	Marrero and Gani, 2001;	$f(GC)$	3.18
Octanol/Water Partition Coefficient Log_{KOW}	Marrero and Gani, 2001;	$f(GC)$	3.18
Water Solubility Coefficient, Log_{WS}	Marrero and Gani, 2001;	$f(GC)$	3.18
Acid Dissociation Constant, pKa^{298K}	Marrero and Gani, 2001;	$f(GC)$	3.18
Auto Ignition Temperature, AiT	Marrero and Gani, 2001;	$f(GC)$	3.18
Flash point, T_f	Marrero and Gani, 2001;	$f(GC)$	3.18
Liquid Viscosity at 298 K, $Visc_L^{298K}$	Marrero and Gani, 2001;	$f(GC)$	3.18
Liquid thermal conductivity, $Therm.COND_L^{298K}$	Marrero and Gani, 2001;	$f(GC)$	3.18
Fathead Minnow 96-hr, $-\log LC_{50}^{FM}$	Hukkerikar et al., 2012a	$f(GC)$	3.18
Daphnia Magna 48-hr, $-\log LC_{50}^{DM}$	Hukkerikar et al., 2012a	$f(GC)$	3.18
Oral Rat LD50, $-\log LD_{50}$	Hukkerikar et al., 2012a	$f(GC)$	3.18
Bio-concentration factor, $logBCF$	Marrero and Gani, 2001;	$f(GC)$	3.18
Permissible exposure limit (OSHA-TWA), $-$logPEL	Marrero and Gani, 2001;	$f(GC)$	3.18
Photochemical oxidation potential, PCO	Marrero and Gani, 2001;	$f(GC)$	3.18
Global warming potential, GWP	Marrero and Gani, 2001;	$f(GC)$	3.18
Ozone depletion potential, ODP	Marrero and Gani, 2001;	$f(GC)$	3.18

Table 3.12. Property model of pure compounds – primary properties (3)

Property	Model	Function	Eq.
Acidification potential, AP	Marrero and Gani, 2001a;	$f(GC)$	3.18
Emission to Urban Air (Carcinogenic), $-\log(EUAC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Urban Air (Non-Carcinogenic), $-\log(EUANonC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Rural Air (Carcinogenic), $-\log(ERAC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Rural Air (Non-Carcinogenic), $-\log(ERANonC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Urban Air (Carcinogenic), $-\log(EUAC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Urban Air (Non-Carcinogenic), $-\log(EUANonC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Rural Air (Carcinogenic), $-\log(ERAC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Rural Air (Non-Carcinogenic), $-\log(ERANonC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Fresh Water (Carcinogenic), $-\log(EFWC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Fresh Water (Non-Carcinogenic), $-\log(EFWNonC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Sea Water (Carcinogenic), $-\log(ESWC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Sea Water (Non-Carcinogenic), $-\log(ESWNonC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Natural Soil (Carcinogenic), $-\log(EASC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Emission to Agriculrural Soil (Non-Carcinogenic), $-\log(EASNonC)$	Hukkerikar et al., 2012b	$f(GC)$	3.18
Cloud Point, CP	Mattei et al., 2014;	$f(GC)$	3.18
Critical Micelle Concentration, CMC	Mattei et al., 2014;	$f(GC)$	3.18
Hydrophilic-Lipophilic Balance, HLB	Davies et al., 1957;	$f(GC)$	3.20
Higher heating value, HHV	Yunus et al.,2014	$f(GC)$	3.18
CO_2 emission in combustion engine, CO_2E	Inventory of U.S. Greenhouse Gas Emissions and Sinks	$f(V_{fuel}, HHV, C_{content,i},$	3.21

Table 3.13. Property model of pure compounds– secondary properties

Property	Model	Function	Eq.
Critical Compressibility Factor, Z_c	Theoretical Equation (PV/RT)	$f(Pc, Vc, Tc)$	3.22
Entropy of Fusion, ΔS_{fus}	Theoretical Equation ($\Delta H_{fus} * 1000/T_m$)	$f(\Delta H_{fus}, Tm)$	3.23
Liquid Volume at T_b , V_L^{Tb}	Tyn and Calus Correlation, Reid et. al. 1987	$f(Vc)$	3.24
Refractive Index, RD	Reid et. al. 1987	$f(Vc)$	3.25
Molar Refraction, $MolarRefraction$	Correlation	$f(RD, V_m^{298K})$	3.26
Dipolar Moment, DipolarMoment	Wilson, 1996	$f(Sol.Pa, Vm)$	3.27
Dielectric Constant, DielectricConstant	Wilson, 1996;	$f(RD, \delta)$	3.28
Henry Constant at 298 K, $HenryConst^{298K}$	Wilson, 1996	$f(P_{vap}^{298K}, Ws\ , Mw)$	3.29
Cost, C	Conte, et al., 2011	$f(Vm)$	3.30
Krafft Temperature, T_K	Correlation (Li et al, 2007)	$f(Log_{KOW}, \Delta H_{formation})$	3.31

Table 3.14. Property model of pure compounds – functional properties

Property	Model	Function	Eq.
Diffusion coefficient at infinite dilution in water, $Diff.Coeff$	Modified Tyn & Calus Correlation, Reid et.al.1987	$f(T, V_L^{Tb})$	3.32
Liquid Density, ρ	Modified Rackett correlation (Reid et.al.1987); PCSAFT (Gross and Sadowski, 2001)	$f(T, Tc, Pc, \omega)$ $f(T, Mw, m, \sigma, \varepsilon/k)$	3.33
Liquid Thermal Conductivity, k	Reid et. al.1987; Gharageizi et al., 2012	$f(T, Vw, Tc, Tb)$ $f(T, \omega, Pc, Tb, Mw)$	3.34 3.35
Vapor Pressure, P^{sat}	Modified SRK EOS; Ceriani et al., 2013; PCSAFT (Gross and Sadowski, 2001)	$f(T, Tc, Pc, \omega)$ $f(T, Mw, GC)$ $f(T, Mw, m, \sigma, \varepsilon/k)$	3.36
Enthalpy of Vaporization, ΔH_{vap}	Reid et. al. 1987; Ceriani et al., 2013; PCSAFT (Gross and Sadowski, 2001)	$f(T, Tc, \omega)$ $f(T, Mw, GC)$ $f(T, Mw, m, \sigma, \varepsilon/k)$	3.37 3.38
Hildebrand Solubility Parameter, δ	Theoretical Equation;	$f(T, \Delta Hvap, Vm)$	3.39
Ideal Gas Heat Capacity, C_p^{ideal}	Reid et. al.1987;	$f(T, GC)$	3.40
Liquid Heat Capacity, C_p^{liquid}	Reid et. al.1987; Diaz Tovar et al., 2011	$f(T, Tc, C_p^{ideal}, \omega)$ $f(T, GC)$	3.41 3.42
Liquid Surface Tension, σ	Diaz Tovar et al., 2011; Brock and Bird equation (Reid et. al.1987); Mattei et al., 2014	$f(T, GC)$ $f(T, Tc, Pc)$ $f(T, GC)$	3.43 3.44 3.45

Vapor Viscosity, η^{gas}	Reid et. al. 1987;	$f(T, Tc, \text{DipolarMoment})$	3.46
Vapor Thermal Conductivity, λ	Eucken and Modified Eucken Model (Reid et. al.1987)	$f(T, Cp, Mw, \eta^{gas})$	3.47
Vapor Volume (V^{gas})	PCSAFT (Gross and Sadowski, 2001)	$f(T, Mw, m, \sigma, \varepsilon/k)$	
Evaporation Time, T90	Conte et al., 2011	$f(\Delta P_{vap})$	3.48
Dynamic Liquid Viscosity, η	Reid et. al.1987;	$f(T, Mw, GC)$	3.49
	Ceriani et al., 2011;	$f(T, Mw, GC)$	3.50
	* This work	$f(T, GC)$	

Table 3.15. Property model equations – primary properties

Eq.	Model	Parameter	Information
3.18	$f(X) = \sum_{i}^{NG1} N_i C_i + w \sum_{i}^{NG2} M_j D_j + z \sum_{i}^{NG3} O_k E_k$	C_i is the contribution for the first-order group of type i with N_i occurrences; D_j is the contribution for the second-order group of type j with M_j and E_k is the contribution of the third-order group of type k with O_k occurrences; and w and z are the constants for the second-order and third-order groups, respectively	These group contribution methods of Marrero and Gani (2001), Hukkerikar et al (2012a), Hukkerikar et al (2012b), Mattei et al. (2014) and Yunus et al. (2014) are determined through three three-step regression procedure
3.19	$tr = 1.0 - 298.15/Tc$ $Zra = 0.29056 - 0.08775 \cdot \omega$ $temp = 1 + (1 - tr)^{0.28571}$ $Vm = (83.14 \cdot Tc \cdot Zra^{temp})/Pc$	Tc (K) is critical temperature; Pc (bar) is critical pressure; ω is acentric factor	
3.20	$HLB = 7 + (n_{i,h} c_{i,h}) - (n_{i,l} c_{i,l})$	n is the number of groups of type i in the molecule, and C is the respective contri-bution. The sub-scripts h and l, instead, refer to the hydrophilic and lipophilic groups, respectively	HLB has been proposed by Davies (1957) and it is a group-contribution based model
3.21	$CO_2\text{E} = (\frac{44}{12}) \sum_{i=1}^{n} V_i \cdot HHVi \cdot C_{content,i} \cdot FO_i$	CO_2E is carbondioxide emission; V_i is volume of fuel type i; $HHVi$ is higher heating value of fuel type i; $C_{content,i}$ is carbon content coefficient of fuel type i; FO_i is fraction oxidized of fuel type i	

Table 3.16. Property model equations – secondary properties

Eq.	Model	Parameter	Information
3.22	$Z_c = P_c V_c / 83.14 T_c$	P_c (bar) is the critical pressure; V_c (cm^3/mol) is the critical volume; and T_c (K) is the critical temperature	These group contribution methods of Marrero and Gani (2001), Hukkerikar et al (2012a), Hukkerikar et al (2012b), Mattei et al. (2014) and Yunus et al. (2014) are determined through three three-step regression procedure
3.23	$\Delta S_{fus} = 1000 \Delta H_{fus} / T_m$	ΔS_{fus} is entropy of fusion (J/mol·K); ΔH_{fus} (kJ/mol) is the enthalpy of fusion; and T_m is the normal melting point	
3.24	$V_L^{Tb} = 0.285 Vc^{1.048}$	V_L^{Tb} (cm^3/mol)is liquid volume; V_c (cm^3/mol) is the critical volume	
3.25	$RD = (0.48872\delta^{298K} + 5.55)/9.55$	RD is reflactive index; δ^{298K}($MPa^{1/2}$) is the solubility parameter at 298 K	
3.26	$MolarRefraction = (1000 V_m^{298K} \cdot (RD^2 - 1))/(RD^2 + 2)$	Molar refraction (cm^3/mol); V_m^{298K}(cm^3/mol) is liquid volume at 298 K	
3.27	$DipolarMoment = 0.0267 \delta_p \cdot V_m^{298K\,5}$	Dipolar moment (debye); V_m^{298K}(cm^3/mol) is liquid volume at 298 K; δ_p is the Hansen polar solubility ($MPa^{1/2}$)	
3.28	when dipolar moment is very near to zero, $DielectricConstant = RD^2$ when dipolar moment is not near zero, $DielectricConstant = (0.48871\delta - 7.5)/0.22$	RD is refractive index; δ_p is solubility parameter ($MPa^{1/2}$)	
3.29	$HenryConst^{298K} = Mw P_{sat}^{298K} / Ws$	Mw (g/mol) is molecular weight; P_{sat}^{298K} is vapor pressure at 298 K; and Ws (mg/L) is water solubility.	
3.30	$C = 2.152 V_m - 38.714$ $C = 2.356 V_m - 119.00$	C is the cost ($/kmol); and *Vm* is the molar volume (l/kmol). It is necessary to notice, however, that this model is not accurate, as it does not take into account the fluctuations of the market, and it has been developed only for preliminary selection purposes, when the cost of several potential candidates cannot be retrieved	Conte et al. (2011) proposed a simple correlation in order to provide qualitatively correct estimations of the pure compound cost, as a function of the molar volume. The first equation is for alcohols and the second equation is for esters.
3.31	$T_K = 57.4 - 7.6 KS2 - 0.06 \Delta H_{formation} + 47.1 A Log_{KOW} - 28 A Log_{KOW} - 36.1 IC + 6.7 nO$	Tk (K) is Krafft temperature; $\Delta H_{formation}$ (Heat of formation); Log_{KOW} is octanol-water partition coefficient; KS2, A, P98, IC and nO are the model descriptors	

Table 3.17. Property model equations – functional properties (1)

Eq.	Model	Parameter	Information
3.32	$X = \exp(-24.71 + \frac{4209}{T} + 0.04527T - 0.00003376T^2)$ $Diff.Coeff = 0.01955X/(Vb^{0.433}T)$	$Diff.Coeff$ (cm^3/s) is infinite dilution in water (cm^3/s) ; V_L^{Tb} (cm^3/mol)is liquid molar volume at normal boiling point temperature; and T (K) is temperature	
3.33	$Zra = 0.29056 - 0.08775\ \omega$ $Temp = 1 + (1 - \frac{T}{Tc})^{0.285714}$ $\rho = (83.14\ Tc \cdot Zra^{Temp})/Pc$	ρ (g/cm^3) is liquid density; Tc (K) is critical temperature, Pc (bar) is critical pressure; ω is Pitzer's Acentric Factor; and T is temperature (K).	Liquid density can be calculated using modified Rackett correlation as shown in these equations. ρ also can be calculated from V_L which can be calculated by employing PCSAFT algorithms in VPPD-Lab (Gross and Sadowski, 2001). The input parameters for the calculation are: Mw (molecular weight, g/mol); m (segment number); σ (segment diameter, Å); ε/k is segment energy parameter (K); and T (K) is temperature.
3.34	$Tr = T/Tc$ $Tbr = \frac{Tb}{Tc}$ $k = [1.11/(Mw^{\frac{1}{2}}) \cdot (3 + 20 \cdot (1 - Tr)^{0.6666})]/[(3 + 20 \cdot (1 - Tbr)^{0.6666})]$	k (W/m·K) is liquid thermal conductivity; Mw (g/mol) is molecular weight; Tc (K) is critical temperature; Tb (K) is normal boiling point; and T (K) is temperature	
3.35	$k = 10^{-4}(10\omega + 2Pc - 2T + 4 + 1.908(Tb + \frac{1.009B^2}{Mw^2}) + \frac{3.9287Mw^4}{B^4} + \frac{A}{B^8})$ $A = 3.8588Mw^8(1.0045B + 6.5152Mw - 8.9756)$ $B = 16.0407Mw + 2Tb - 27.9074$	k (W/m·K) is liquid thermal conductivity; Tb (K) is normal boiling point; ω is centric factor; Pc (bar) is critical pressure; Mw (g/mol) is molecular weight; and T (K) is temperature	For lipid compounds such as fatty acids, alcohols, esters and acylglycerols, liquid thermal conductivity can be calculated by the model from Gharageizi et al. (2012)
3.36	$ln(P^{sat}) = A + B/T + C * ln(T))$ $A = \sum(Nk \cdot (A1k + Mw \cdot A2k)) + (so + Ncs * s1) + Alfa * (fo + Nc * f1)$ $B = \sum(Nk \cdot (B1k + Mw \cdot B2k)) + Beta * (fo + Nc * f1)$ $C = \sum(Nk \cdot (C1k + Mw \cdot C2k))$	P^{sat} (Pa) is vapor pressure; Nk is the number of groups k in the molecule; Mw is the molecular weight of the component; Ncs is the number of carbon atoms in the alcohol part of esters; Nc is the total number of carbon atoms; A1k, B1k, C1k, A2k, B2k, C2k, Alfa, Beta, so, s1, fo, f1 are model parameters; and T (K) is temperature	Vapor pressure is able to be calculated by employing PCSAFT algorithms in VPPD-Lab (Gross and Sadowski, 2001) as well as modified SRK equation of state (SRK EOS). The input parameters for PCSAFT calculation are: Mw (molecular weight, g/mol); m (segment number); σ (segment diameter, Å); ε/k is segment energy parameter (K); and T (K) is temperature. For lipid compounds such as fatty acids, alcohols, esters and acylglycerols, vapor pressure can be calculated by these equations (Ceriani et al., 2013)

Table 3.18. Property model equations – functional properties (2)

Eq.	Model	Parameter	Information
3.37	$Tr = 1 - \mathrm{T/Tc}$ $W = (\omega - 0.21)/0.25$ $R1 = 6.537\mathrm{tr}^{0.333} - 2.467\mathrm{tr}^{0.833} - 77.521\mathrm{tr}^{1.208} + 59.634\mathrm{tr} + 36.009\mathrm{tr}^2 - 14.606\mathrm{tr}^3$ $R2 = 0.133\mathrm{tr}^{0.333} - 28.215\mathrm{tr}^{0.833} - 82.958\mathrm{tr}^{1.208} + 99\mathrm{tr} + 19.105\mathrm{tr}^2 - 2.796\mathrm{tr}^3$ $\Delta H_{vap} = (\mathrm{R1} + \mathrm{W} \cdot \mathrm{R2}) \cdot \mathrm{Tc} \cdot 0.008314$	ΔH_{vap} (kJ/mol) is enthalpy of vaporization, Tc (K) is critical temperature; ω is acentric factor; and T (K) is temperature	Reid et al. (1987)
3.38	$\Delta H_{vap} = 8.3144 \cdot (-\mathrm{B} + \mathrm{C} \cdot \mathrm{T}) \cdot \left(1 - \left(\frac{\mathrm{T}}{\mathrm{Tc}}\right)\right)$ $B = \sum(\mathrm{Nk} \cdot (\mathrm{B1k} + \mathrm{Mw} \cdot \mathrm{B2k})) + \mathrm{Beta} \cdot (\mathrm{fo} + \mathrm{Nc} \cdot \mathrm{f1})$ $C = \sum(\mathrm{Nk} \cdot (\mathrm{C1k} + \mathrm{Mw} \cdot \mathrm{C2k}))$	ΔH_{vap} (J/gmol) is enthalpy of vaporization, Tc (K) is critical temperature; Nk is the number of groups k in the molecule; Mw is the molecular weight of the component; Ncs is the number of carbon atoms in the alcohol part of esters, Nc is the total number of carbon atoms; B1k, C1k, B2k, C2k, Alfa, Beta, fo, f1 are model parameters; and T (K) is temperature	The correlation for calculation of enthalpy of vaporization is available in Ceriani et al. (2013)
3.39	$\delta = \left[\frac{1000 * \Delta H_{vap}(\mathrm{T}) - 8.314 \cdot \mathrm{T}}{V_m(\mathrm{T})}\right]^{1/2}$	δ ($\mathrm{MPa}^{1/2}$) is Hildebrand Solubility Parameter; ΔH_{vap} (kJ/mol) is enthalpy of vaporization; V_m(cm^3/mol) is liquid volume; and T (K) is temperature	
3.40	$C_p^{ideal} = \sum N_k (A_k + (B_k T) + (C_k T^2) + (D_k T^3))$	C_p^{ideal} $(J/mol \cdot K)$ is ideal gas heat capacity; Nk is the number of groups k in the molecule; Ak, Bk, Ck, Dk are model parameters; and T (K) is temperature	
3.41	$tr = T/Tc$ $temp = 1.45 + \frac{0.45}{1 - tr} + 0.25 \cdot omega \cdot \left[17.11 + \frac{25.2}{Tr} \cdot 1 - tr)^{0.33333} + \frac{1.742}{1 - tr}\right]$ $C_p^{liquid} = 8.314 \cdot temp + C_p^{ideal}$	C_p^{liquid} $(J/mol \cdot K)$ is liquid heat capacity; Tc (K) is critical temperature, ω is acentric factor; C_p^{ideal} $(J/mol \cdot K)$ is ideal gas heat capacity; and T (K) is temperature	
3.42	$C_p^{liquid} = \sum N_k (A_k + (B_k T))$	C_p^{liquid} $(J/mol \cdot K)$ is liquid heat capacity; Nk is the number of groups k in the molecule; Ak, Bk are model parameters; and T (K) is temperature	
3.43	$\sigma = \sum N_k (A1_k + (B1_k T)) + \sum N_k (A2_k + (B2_k T)) + Q$	σ (dynes/cm) is surface tension; N_k is number of groups k in the molecule; Mw is molecular weight of the compound; A1k, B1k, A2k, B2k, are model parameters; Q is the correction terms calculated based on the class of the lipid compound	

Table 3.19. Property model equations – functional properties (3)

Eq.	Model	Parameter	Information
3.44	$Q = 0.1196\left(1+\frac{Tbr\ln\left(\frac{Pc}{1.01325}\right)}{1-Tbr}\right)-0.279$ $\sigma = P_c^{\frac{2}{3}}T_c^{\frac{1}{3}}Q(1-T_r)^{11/9}$	σ is surface tension; Tbr is reduced temperature at the normal boiling point; Pc (bar) is critical pressure; Tc (K) is critical temperature	
3.45	σ=11.98+0.478·nO+0.5848·KHO−0.0007763·ET−0.0103· $\Delta Hform$+0.09734·D−0.1345·nO·KHO	σ is surface tension; nO, KHO, ET and D are the model descriptors	
3.46	$\eta\S = [0.807T_r^{0.618} - 0.357\exp(-0.449T_r) + (0.34\exp(-4.058T_r) + 0.018]F_p^0F_Q^0$ $\S = 0.176(\frac{T_c}{P_c^4Mw^3})$ $\mu_r = 52.46\left(\frac{\mu^2P_c}{T_c^2}\right)$ $F_p^0 = 1 \quad 0 \le \mu_r \le 0.022$ $F_p^0 = 1+30.55(0.292-Z_c)^{1.72}$ $0.022 \le \mu_r \le 0.075$	η (μP) is gas viscosity; § is energy-potential parameter; T_c(K) critical tempature; T_r, reduced temperature; P_c (bar) is critical pressure; Mw (g/mol) is molecular weight; F_p^0 is low pressure polar correction factor; $F_Q^0$is low pressure quantum correction factor; μ (debyes) is dipole moment; μ_r is dimensionless dipole moment; Z is compressiblity factor; Z_c is critical compressibility factor; and T (K) is temperature	F_Q^0 is used only for the quantum gases such as H2, H2, and D2.
3.47	$\frac{\lambda Mw}{\eta C_v} = 1.32+(\frac{1.77}{C_v/R})$ $C_v = Cp-8.314$	λ (W/m·K) is thermal conductivity; η (μP) is gas viscosity; Mw is molecular weight; C_v $(J/mol\cdot K)$ is specific heat at constant volume; C_p is specific heat at constant pressure	
3.48	$\ln(T90) = -0.793lnP^{sat}+12.416$	$T90$ (s) is evaporation time; and P^{sat} (Pa) is vapor pressure at temperature T (K)	
3.49	$\eta = Mw\cdot 1000\cdot exp(\sum(\mathrm{Nk}\cdot(\mathrm{Ak/T}+\mathrm{Mw}\cdot\mathrm{Bk}))$	η (cp) is dynamic liquid viscosity; Mw (g/mol) is molecular weight; Nk is is the number of groups k in the molecule; Ak, Bk are model parameters; and T (K) is temperature	
3.50	$\ln(\eta) = \sum Nk\cdot(A1k+\frac{B1k}{C1k+T})+\sum Nk\cdot Mw\cdot(A2k+\frac{B2k}{C2k+T})+Q$	η (cp) is dynamic liquid viscosity; Mw (g/mol) is molecular weight; Nk is is the number of groups k in the molecule; A1k, B1k, C1k, A2k, B2k, C2k are model parameters; Qis the correction terms calculated based on the class of the lipid compound such as fatty acids, alcohols, esters and acylglycerols; and T (K) is temperature	For lipid compounds such as fatty acids, alcohols, esters and acylglycerols, dynamic liquid viscosity can be calculated from models from Ceriani et al. (2011)

Some functional properties such as liquid density, liquid thermal conductivity, vapor pressure, enthalpy of vaporization, dynamic viscosity, and liquid surface tension, can be estimated using regression models, where their coefficients were obtained from VPPD-Lab database (Nielsen et al., 2001) as given in **Table 3.20**.

Table 3.20. Correlation equations for functional properties

Correlation	Model equation	Eq.
Solid Density [kmol/m^3]	$A+B\cdot T+C\cdot T^2+D\cdot T^3+E\cdot T^4$	3.51
Liquid Density [kmol/m^3]	$A/B^{(1+(1-T/C)^D)}$	3.52
Vapour Pressure [Pa]	$\exp(A+B/T+C\cdot\ln(T)+D\cdot T^E)$	3.53
Heat of Vaporization [J/kmol]	$A\cdot (1-Tr)^{(B+C\cdot Tr+D\cdot Tr^2)}$	3.54
Solid Heat Capacity [J/(kmol·K)]	$A+B\cdot T+C\cdot T^2+D\cdot T^3+E\cdot T^4$	3.55
Liquid Heat Capacity [J/(kmol·K)]	$A+B\cdot T+C\cdot T^2+D\cdot T^3+E\cdot T^4$	3.56
Ideal Gas Heat Capacity [J/(kmol·K)]	$A+B\cdot (C/T/sinh(C/T))^2+D\cdot (E/T/cosh(E/T))^2$	3.57
Second Virial Coefficient [m^3/kmol]	$A+B/T+(\frac{C}{T})^3+(\frac{D}{T})^8+(\frac{E}{T})^9$	3.58
Liquid Viscosity [kg/(m·s)]	$\exp(A+B/T+C\cdot\ln(T)+D\cdot T^E)$	3.59
Vapour Viscosity [kg/(m·s)]	$A\cdot T^B/(1+C/T+D/T^2)$	3.60
Liquid Thermal Conductivity [J/(m·s·K)]	$A+B\cdot T+C\cdot T^2+D\cdot T^3+E\cdot^{T4}$	3.61
Vapour Thermal Conductivity [J/(m·s·K)]	$A\cdot T^{B/(1+C/T+D/T^2)}$	3.62
Surface Tension [kg/s^2]	$A+B\cdot T+C\cdot T^2+D\cdot T^3+E\cdot T^4$	3.63

Mixture property models

Mixture property models are divided into two classes: linear mixing rule and non-linear mixing rule. The property models and references are given in **Table 3.21**. The property model equations and equation parameters are given in **Table 3.22**– **3.23**.

Table 3.21. Linear and non-linear models for mixture properties

Property	Model	Eq.
Linear mixing rule		3.64
Non-linear mixing rule		
Vapor pressure	Modified Raoult's law	3.65
Flash point	Liaw et al., 2011	3.66
Density	Spence and Danner, 1973	3.67
Dynamic liquid viscosity	Cao et al., 1993	3.68
	Mattei et al., 2014	3.69
The hydrophilic-lipophilic deviation	Salager, 1996	3.70
Surface tension	Wang et al., 2002	3.71
	Suarez et al., 1989	3.72
Distillation curve	Hoffman, 1969	3.73
Hydrophilic-lipophilic balance	Davies, 1957	3.74

Table 3.22. Linear and non-linear model information (1)

Eq.	Model	Parameter	Information
3.64	$P = \sum_{i=1}^{n} x_i P_i$	P_i is the property of component i; x_i is the mass, volume or molar fraction of component i; and n is the number of compounds in mixture.	
3.65	$RVP = \sum_{i=1}^{n} x_i \gamma_i P_i^{sat}$	P_i^{sat} is the saturated vapor pressure at temperature T; and $\gamma_i$is the activity coefficient	The vapor pressure for blended gasoline is referred as the Reid vapor pressure (RVP), which is defined as the vapor pressure measured at a temperature of 100°F (308 K) in a chamber with a vapor/liquid volume ratio of 4:1.
3.66	$\sum_{i=1}^{n} \frac{x_i \gamma_i P_i^{sat}(T)}{P_{i,Tf}^{sat}} - 1 = 0$	P_i^{sat} is the saturated vapor pressure at temperature T; $\gamma_i$is the activity coefficient; $P_{i,Tf}^{sat}$ is vapor pressure of pure compound s at their flash point. The temperature, T is deemed to be the flash point of the mixture. This property model requires an iteration to obtain the flash point of the mixture	Flash point (T_f) is defined as the lowest temperature at which the vapor above a liquid can be ignited in air
3.67	$\frac{1}{\rho_B} = V_{Cm} Z_{RAm}^{\left[1+(1-T_r)^{\frac{2}{7}}\right]}$ $V_{Cm} = \mathrm{R} \sum_{\mathrm{i=1}}^{\mathrm{n}} x_i \frac{T_{ci}}{P_{ci}}$ $Z_{RAm} = \sum_{\mathrm{i=1}}^{\mathrm{n}} x_i Z_{RAi}$ $T_r = \frac{T}{\sum x_i T_{ci}}$	V_{cm} and Z_{RAm} are molar averages of the pure component critical volumes and critical compressibility factors; Z_{RAi} is the particular constant for the Rackett equation for compound *i*. However, it can be replaced with critical compressibility factor, Z_c if it is not available; the unit of measure for mixture density is (mol/cm^3), depending on the universal gas constant, *R*	The modified Rackett equation gives the best prediction of the pure component density for hydrocarbons, and provides a good estimation for organic as well as inorganic compounds. Therefore, the modified Rackett equation was extended for estimation of the mixture's density
3.68	$ln(\eta V) = \sum_{i}^{NC} \varphi_i ln(\eta_i V_i) + 2\sum_{i}^{NC} \varphi_i ln\left(\frac{x_i}{\varphi_i}\right) - \sum_{i}^{NC}\left(\frac{q_i np_i \varphi_i}{r_i}\right)\sum_{j}^{NC} \theta_{ji} ln(\tau_{ji})$ $r_i = \sum_{i}^{NC} \upsilon_{k,i} R_k \quad q_i = \sum_{i}^{NC} \upsilon_{k,i} Q_k$ $\tau_{ij} = exp\left(-\frac{a_{mn}}{T}\right) \varphi_i = \frac{x_i r_i}{\sum_j^{NC} x_j r_j}$ $\theta_{ji} = \frac{\theta_j \tau_{ji}}{\sum_l^{NC} \theta_l \tau_{li}} \quad \theta_j = \frac{x_j q_j}{\sum_i^{NC} x_i q_i}$	η(mPa.s) is the mixture viscosity; V (cm^3/mol) is the mixture volume that can be calculated by **Eq. (3.83).** V_i (cm^3/mol) and η_i (mPa.s) are pure compound molar volume and viscosity;$\upsilon_{k,i}$, R_k and Q_k are group parameters obtained from Magnussen et al (1981); τ_{ij} is calculated from the group interaction parameters a_{mn}; φ is volume fraction; θ_{ji} is parameter;θ_j is surface fraction;	
3.69	$\eta = \eta_s(1 + \varphi\left(\frac{1 + 2.5M}{1 + M}\right))$	η_s is the dynamic viscosity of the continuous phase; M is ratio between the dynamic viscosity of the dispersed phase and that of the continuous phase; and ψ is the volume fraction of the dispersed solvent phase	For emulsions

Table 3.23. Linear and non-linear model information (2)

Eq.	Model	Parameter	Information
3.70	$HLD = \alpha - EON + b \cdot S - K \cdot ACN - \varphi(A) + cT\Delta T$ $HLD = \sigma + ln(S) - K \cdot ACN - f(A) + aT\Delta T$	α, EON and σ are model parameters that can be related; thanks to group-contribu-tion like correlations, to the molecular structure of the surfactants; S is the salinity; ECN is effective carbon number and it is related to the molecular structure of the organic solvent phase, A refers to the alcohol concentration and aT and cT are model parameters defining the temperature dependence.	The hydrophilic-lipophilic deviation (HLD) indicates the stability of an emulsified formulated product, with respect to both temperature and composition disturbances. Negative HLD values suggest the formation of oil-in-water emulsions, positive values suggest the formation of water-in-oil emulsions.Ttwo different correlations have been proposed for non-ionic and ionic surfactants respectively.
3.71	$\sigma = 11.98 + 0.478 \cdot nO + 0.5848 \cdot KHO - 0.0007763 \cdot ET - 0.01053 \cdot \Delta Hform + 0.09734 \cdot D - 0.1345 \cdot nO \cdot KHO$	nO, KHO, ET and D are the model descriptors	For emulsified formulate products. the aqueous phase is usually defined as the surface tension of the aqueous surfactant solution at the critical micelle concentration. The predictive methods can be applied, and the QSPR model proposed by Wang et al (2002) is used.
3.72	$\sigma = \sigma_i + \frac{RT}{A_i}\ln(\frac{x_{i,s}\gamma_{i,s}}{x_{i,b}\gamma_{i,b}})$	σ (mN/m) is surface tension of a mixture; σ_i (mN/m) is the surface tension of the pure compound i; A_i is molar surface area of pure compound i; $x_{i,s}$ and $x_{i,b}$ are composition of compound i in the surface and bulk liquid phase respectively; $\gamma_{i,s}$ and $\gamma_{i,b}$ are activity coefficients of compound i in the surface and bulk liquid phase respectively.	For homogeneous formulated products. It is based on the assumption that the surface layer can be treated as a separate phase located between the vapor and the bulk phases.
3.73	$Y = \ln\left(\frac{T + T_0}{T_0}\right)$ $X = \ln\ln(\frac{1}{1-x_i})$ $XB_T = \frac{1}{C2}$ $A_T = B_T \exp(C_1 B_T)$ $Y = C1 + C2X$	$T_0(K)$ is initial boiling point; T_i (K) is temeperature at which i percent Is distillated; xi is volume or weight part of distillate.	
3.74	$HLB = 7 + n_{i,h} \cdot C_{i,h} - n_{i,l} \cdot C_{i,l}$	n is number of groups of types i in the molecule; C is respective contribution; the sub script h and i refere to the hydrophilic and lipophilic groups respectively	For emulsified formulate products.

Product performance models

The design, development and reliability of a chemical product and the process to manufacture it, need to be consistent with the end-use characteristics of the desired product. One of the common ways to match the desired product-process characteristics is through trial and error based experiments which can be expensive and time consuming. An alternative approach is the use of mathematical models to calculate product performance or product behaviors, replacing some of the time consuming and/or repetitive experimental steps. Furthermore, for many chemical products the appropriate models for product-process design need to have multiscale features as the properties defining the chemical structure and the product end-use characteristics are dependent on parameters of different size and time scales.

Stability check algorithm

This algorithm is used to check the phase stability of a binary liquid mixture. The stability check is based on the trend of Gibbs energy function of mixing ($\Delta G^{mix}/RT$) and its first and second derivatives as function of composition. The Gibbs energy of mixing is calculated as follows:

$$\frac{\Delta G^{mix}}{RT} = \frac{G^E}{RT} + \sum_i^{NC} x_i \ln(x_i) \tag{3.75}$$

G^E is the excess Gibbs energy of mixing that is calculated from:

$$\frac{G^E}{RT} = \sum_i^{NC} x_i \ln(\gamma_i) \tag{3.76}$$

Figure 3.4 represents the four most common type of plots of $\frac{\Delta G^{mix}}{RT}$ as a function of x_i:

(1) Mixture of type a are completely immiscible in the composition range [0,1], and they can be recognized from the positive value of the function $\frac{\Delta G^{mix}}{RT}$ in the entire composition range [0,1];
(2) Mixtures of type b_1 show a phase split in the composition range where the function $\frac{\Delta G^{mix}}{RT}$ is positive; the two liquid phase region corresponds to the region in which the $\frac{\Delta G^{mix}}{RT}$ is positive. The compositions of the two liquid phases are identified by the points in which the function $\frac{\Delta G^{mix}}{RT}$ is zero, at the extremities of the immiscibility gab;
(3) Mixtures of type c are one phase in the entire composition range and they have negative values of the function $\frac{\Delta G^{mix}}{RT}$ and positive values of its second derivative in entire the composition range;
(4) Mixtures of type b_2 are more complex: here,$\frac{\Delta G^{mix}}{RT}$ is negative and its second order derivative is negative between the compositions x_1^δ and x_2^δ (δ and ε are the wrong phases in equilibrium). These mixtures show a miscibility gap, but the compositions of the two liquid phases are not identified by the points in which the second derivative of $\frac{\Delta G^{mix}}{RT}$ changes its sign (x_1^δ and x_2^δ), since these points do not correspond to the composition at which the total Gibbs energy is at its global minimum. In order to identify the composition of the two liquid phases in equilibrium, the tangent plane condition has to be employed.

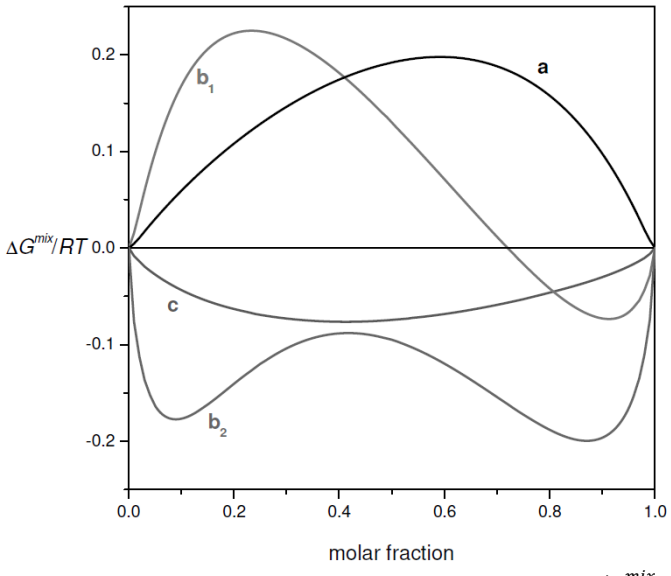


Figure 3.4. The most common shapes for the function $\frac{\Delta G^{mix}}{RT}$

This algorihm was developed by Conte et al (2011). The algorithm consists of three levels of screening. The UNiFAC model with the LLE group decomposition and contributions (Magnussen et al., 1981) has been chosen to describe the binary chemical systems. The input of this algorithm is the UNIFAC-LLE group decomposition of the compounds involved in the mixtures under evaluation, and the temperature at which the stability test has to be performed. The algorithm returns, as output, the stability information about the solvent mixtures (total miscibility, partial miscibility, total immiscibility), and , in case of partially miscible mixtures, the compositions of the two phases in equilibrium is also given.

Microcapsule for controlled release of active ingredients

Microcapsules are the products of the microencapsulation that basically consists of the wrapping of substances (one playing the role of active ingredient and the other solvent medium) in another substance (shell of the microcapsule). Microcapsules usually have a spherical shape on a small size scale, where the wall/shell of the microcapsule regularly envelops the core that consists of the active ingredient and solvent. In some cases, however, the droplets of the active ingredient plus solvent medium are dispersed in the complete body of the microcapsule, where the phase of the active ingredient and/or solvent might be liquid or gas.

In the agrochemical industry, the use of controlled release technology for the delivery of pesticides to the environment has numerous advantages: from optimized delivery of the active ingredients, to reduction of hazards to humans and environment. That is, the amount of pesticide used on the field can be reduced and also the safety level during its use can be improved. As far as controlled release of drugs or pharmaceutical products is

concerned, this technology is mainly aimed at controlling the amount of the drug delivered. The main benefits of controlled release are that it helps to keep an effective level of drug in the body for a specific period of time, and thereby, side effects generated by drug overdosing and/or under-dosing may be avoided.

Figure 3.5 shows the schematic representation of one microcapsule, where the active ingredient (AI, with concentration C_d) is enclosed within the core of the device inside radius r_i) , in a polymeric membrane of thickness $r_0 - r_i$. The AI from the core of the microcapsule is delivered to the release medium with concentration C_r. Thus, the performance of the microcapsule for controlled release can be predicted through the use of a mathematical model.

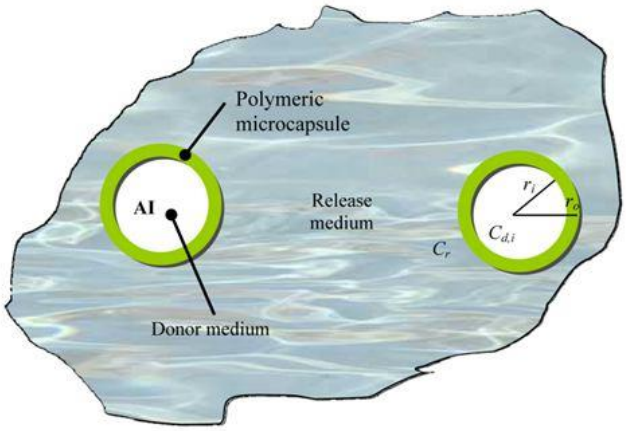


Figure 3.5. Schematic representation of a microcapsule

Brief descriptions of the main items within the system are given below:

- Active ingredient (AI): This is the substance that is enclosed in the core of the microcapsule. The identity of the substance differs according to the applications of the microcapsule device. For instance, if the microcapsule is applied for pest control, a pesticide or fungicide is in the core of the microcapsule. For delivery of pharmaceutical product, the substance is a drug such as, antibiotic, antibodies, antioxidants, probiotics, etc., placed in the core of the device;
- Donor medium: Usually, the active ingredients are solid substances that need to be dissolved in a solvent (donor medium);
- Microcapsule wall (polymeric membrane): The active ingredient together with the solvent is encapsulated by a polymeric membrane. One of the most important phenomena occurring here is the mass transfer of the active ingredient through this wall;
- Release Medium: The active ingredient diffuses out from the microcapsules into the release medium. The medium depends of the application field; for instance,

in the agrochemical field the most common release medium is water, while for a pharmaceutical product, the release medium could be gastric acid, blood, or any other medium found within the human body.

A mathematical model (Muro-Suñé et al., 2005) representing the controlled release of Pesticide delivery to the environment considers the following three different scales: nano-scale (calculation of diffusion coefficient), micro-scale (release of active ingredient in presented microcapsules), and meso-scale (normal distribution of microcapsules).

For nano-scale, the diffusion coefficient is calculated as follows:

$$D_1 = D_0 \exp\left(\frac{-E}{RT}\right) \exp\left(\frac{-\left(w_1 \widehat{V_1^*} + w_2 \xi \widehat{V_2^*}\right)}{w_1 \frac{K_{11}}{\gamma}\left(K_{21} - T_{g1} + T\right) + w_2 \frac{K_{12}}{\gamma}\left(K_{22} - T_{g2} + T\right)}\right) \quad (3.77)$$

where,

D_1 = diffusion coefficient of AI in the polymer
D_0 = constant pre-exponential factor.
E = energy (per mole) that the molecule needs to overcome attractive forces which constrain it to its neighbors.
R = gas constant.
T = temperature.
T_{gi} = glass transition temperature of compound i.
w_i = weight fraction of component i.
K_{1i}, K_{2i} = free volume parameter of compound i.
$\widehat{V_1^*}$= critical hole free volume required for a jump.
ξ = ratio of molar volumes for the solvent and the polymer jumping units.
γ = overlap factor (between 0.5 and 1).

For micro-scale, equations at the micro-scale represent the behavior of one single microcapsule These mathematical expressions describe the concentration of the active ingredient in the donor medium as well as in the release medium. The concentration profile is following Fick's law of diffusion and has a dependence on the time. The concentration of the active ingredient in the donor is given by:

$$\frac{dC_{d,i}}{dt} = -\frac{DA_i}{hV_{d,i}} K_{\frac{m}{d}} C_{d,initial} \exp\left(-\frac{DA_i K_{\frac{m}{d}}}{V_r h}\left(\frac{K_{\frac{m}{r}}}{K_{\frac{m}{d}}} + \frac{V_r}{V_{d,i}}\right) t\right) \quad (3.78)$$

where,

C = concentration of donor as function of time (g/m^3).
$C_{d,initial}$ = initial concentration of donor as function of time (g/m^3).
$K_{m/d}$ = partition coefficient of the AI between the donor and the polymer membrane.

$K_{m/r}$ = partition coefficient of the AI between the receiver and the polymer membrane.
$V_{d,i}$ = donor volume (m^3).
V_r = receiver volume (m^3).
A_i = surface area through which diffusion takes place (m^2).
D = polymer solvent binary mutual diffusion coefficient (m^2/s).
h = thickness of the microcapsule wall (m).
t = time (s).

Eqs. (3.78) describes the change of the concentration of the active ingredient in the donor with respect to time. Analytical solution of **Eqs. (3.78)** gives the following result:

$$C_{d,i} = \frac{K_{\frac{m}{r}} V_{d,i}}{K_{\frac{m}{d}} V_r + K_{\frac{m}{r}} V_{d,i}} C_{d,initial}(1 - \exp(-\frac{DA_i}{V_r h} K_{\frac{m}{d}} (\frac{K_{\frac{m}{r}}}{K_{\frac{m}{d}}} + \frac{V_r}{V_{d,i}})t)) \quad (3.79)$$

The concentration of the active ingredient in the receiver (release medium) is given by:

$$\frac{dC_{r,i}}{dt} = \frac{DA_i}{V_r h} K_{\frac{m}{d}} C_{d,initial} \exp(-\frac{DA_i}{V_r h} K_{\frac{m}{d}} (\frac{K_{\frac{m}{r}}}{K_{\frac{m}{d}}} + \frac{V_r}{V_{d,i}})t) \quad (3.80)$$

Eq. (3.80) is based on Fick's law of diffusion, represents the change with respect to time of the active ingredient in the release medium. It is easy to observe that **Eqs. (3.78)** and **(3.80)** are strongly affected by two properties; the diffusion coefficient of the active ingredient in the polymer and the partition coefficient between the polymer membrane, and, the donor and the polymer of membrane and the release medium. The analytical form of **Eqs. (3.80)** is the following:

$$C_{r,i} = \frac{K_{\frac{m}{r}} V_{d,i}}{K_{\frac{m}{d}} V_r + K_{\frac{m}{r}} V_{d,i}} C_{d,initial}(1 - \exp(-\frac{DA_i}{V_r h} K_{\frac{m}{d}} (\frac{K_{\frac{m}{r}}}{K_{\frac{m}{d}}} + \frac{V_r}{V_{d,i}})t) \quad (3.81)$$

Burst and lag time effects

The controlled release model can be further refined in order to predict more accurately the initial periods of release. In this initial period, before an eventual steady-state is reached, it is necessary to account for the so-called burst and lag time effects. These phenomena depend mainly on the diffusivity of the solute in the polymer, the thickness of the membrane and the storage as well as the usage conditions.
The *burst effect* occurs when, for example, the devices are stored for a period, giving time for the AI to diffuse into the polymer membrane and saturate it. Thus, when the system is used, the initial delivery rate from the microcapsule becomes greater than that of the steady state. This is described as follows:

$$M(t)_{r,i} = \frac{V_r C'_{d,initial}}{\left(\frac{K_{\frac{m}{r}}}{K_{\frac{m}{d}}} + \frac{V_r}{V_{d,i}}\right)}(1 - \exp(-\alpha t)) + J_{max} A_i \frac{2}{\alpha'}(1 - \exp(-\alpha' t)) \quad (3.82)$$

where,
$M_{r,i}\ (t)$ = concentration of receiver as a function of time (*g/m*3).
$C'_{d,initial}$ = modified initial concentration of donor as function of time (*g/m*3).

In the *lag effect*, there is no lapse between fabrication and use of the device, the active ingredient does not have time to partition into the membrane and there is a delay before the steady state gradient is reached. The lag profile of the active ingredient can be described as follows:

$$M(t)_{r,i} = \frac{V_r C'_{d,initial}}{\left(\frac{K_{\frac{m}{r}}}{K_{\frac{m}{d}}} + \frac{V_r}{V_{d,i}}\right)}(1 - \exp(-\alpha t)) + J_{max} A_i \frac{2}{\alpha'}(1 - \exp(-\alpha' t)) \quad (3.83)$$

where,

$$\alpha = \frac{D A_i}{V_r h} K_{\frac{m}{d}} \left(\frac{K_{\frac{m}{r}}}{K_{\frac{m}{d}}} + \frac{V_r}{V_{d,i}}\right) \quad (3.84)$$

$$\alpha' = \frac{D\pi^2}{h2} K_{\frac{m}{d}} \quad (3.85)$$

$$J_{max} = \frac{D D_{d,initial}}{h} K_{\frac{m}{d}} \quad (3.86)$$

It is important to highlight that it is considered that the initial concentration (**Eqs. (3.87)**) used in the first-order release terms for burst and lag time effect are:

$$C'_{d,initial,i} = \frac{M'_{d,initial,i}}{V_{d,i}} \quad (3.87)$$

$$M'_{d,initial,i} = M'_{d,initial,i} - M'_{\frac{burst}{lag},\infty,i} \quad (3.88)$$

$$M_{d,initial,i} = C_{d,initial} V_{d,i} N_{p,i} \quad (3.89)$$

For instance, if the product is delivered almost instantaneously the burst effect becomes:

$$M_{\frac{burst}{lag},\infty} = \frac{2}{\alpha'} J_{max} A_i N_{p,i} \quad (3.90)$$

If the release is not carried out initially, a lag time effect is added and the initial concentration is calculated as follows:

$$M_{lag,\infty} = -\frac{2}{\alpha'} J_{max} A_i N_{p,i} \tag{3.91}$$

For meso-scale, it is considered that not all the microcapsules have the same size. Hence, the meso-scale calculations account for the number of microcapsules of different sizes and consider the distribution of sizes of microcapsules. The number of microcapsules and their sizes are considered through a normal distribution function as follows:

$$F(r;\mu;\sigma) = \int_{-\infty}^{r'} \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(r-\mu)^2}{2\sigma^2}\right) dr \tag{3.92}$$

where,

r = microcapsule radius.
μ = mean distribution value.
σ = standard deviation.

The equations representing the mathematical model of the microcapsule based controlled release involve the use of different integration frameworks to link the models at different scales. The connection between micro-scale and macro-scale involves the simultaneous integration strategy. With respect to micro-scale and nano-scale connection, serial-one way coupling integration is used because the values of the diffusion coefficients are calculated using **Eqs. (3.77)** or via dynamic simulation. The connection can also involve a serial-transformation integration structure, if the values of diffusion coefficients are obtained through the use of a correlation obtained by fitting experimental data.

3.3.3.2 Lipids database

Lipids are organic compounds that are insoluble in polar solvents (such as water), and soluble in organic solvents (such as chloroform and acetone) and alcohols. There pure compound and mixture properties are necessary for synthesis, design, and analysis of processes for the production of edible oils, fats, biodiesel, and other lipids. The lack of measured data for these systems makes it necessary to develop reliable predictive models based on limited data. Therefore, SPEED Lipids database has been developed. The database contains information on 331 compounds (see **Appendix A.1**). **Table 3.29** lists compound classes present in the database.

Table 3.24. Classification of compounds in SPEED Lipids database

No.	Class	Number of compounds	Carbon range
1	Carotenoids	8	C_{40}
2	Diglycerides	41	C_{14}-C_{43}
3	Ethyl esters	28	C_{8}-C_{26}
4	Ethylhexyl esters	9	C_{14}-C_{26}
5	Fatty acids	29	C_{6}-C_{24}
6	Isopropyl esters	3	C_{19}-C_{21}
7	Methyl esters	28	C_{7}-C_{25}
8	Monoglycerides	15	C_{11}-C_{25}
9	Pesticides	14	C_{9}-C_{20}
10	Triterpenealcohols	8	C_{8}-C_{30}
11	Others	13	C_{41}-C_{47}
12	Ubiquinones	5	C_{27}-C_{29}
13	Phospholipids	14	C_{22}-C_{41}
14	Sterol-esters	7	C_{41}-C_{47}
15	Sterols	5	C_{27}-C_{29}
16	Triglycerides	85	C_{27}-C_{75}
17	Vitamin E	9	C_{27}-C_{52}
18	Fatty alcohol	10	C_{6}-C_{22}

The database is divided into two parts:

- Pure compound properties: the database contains experimental data 7 primary properties and 5 functional (temperature dependent) properties. The list and number of experimental data are shown in **Appendix A.2** (for primary properties) and **Appendix A.3** (for temperature dependent properties). In order to estimate the missing properties that are needed for model-based design and analysis of edible oil and biodiesel processes, suitable property models for each class of compounds as listed in section 3.3.3 are employed (see **Table 3.30**);
- Mixture properties: the database contains about 4500 measured data points for 332 different phase equilibrium data-sets, including binary and multi-component systems (92 VLE, 91 LLE, 70 SLE and 79 solubility data). Uncertainties of experimental measurements or quality estimates given by the authors of the experimental were also considered (Cunico et al., 2014). The published activity coefficients and parameter values from fitting different G^E models (UNIQUAC, original UNIFAC and NRTL) for VLE and SLE binary systems are stored in the database. These models were used in parameter regressions for fine-tuning existing model parameters, improving VLE and SLE predictions, and obtaining model parameters not available in the literature.

An important issue is to use pure compound properties in order to calculate mixture properties as well as process models is the accuracy and reliability of the property data to be used in the calculation. The accuracy of the experimental data and the estimated data for each class of compounds can be viewed by plotting the property values against increasing carbon number. **Figure 3.6** shows an example of the consistency check between experimental data of boiling point (Tb) and estimated values. The larger molecule tends to have higher Tb because the larger molecule needs more energy to break the cohesive interactions in the liquid phase to become molecule in the gas phase. The boiling point experimental and estimated values of fatty acids are close to each other and follow the boiling point trend. Therefore, boiling point experimental and estimated data of fatty acids are consistent. The consistency test results for all properties in the database are shown in **Appendix A.3**.

Table 3.25. Classification of compounds in SPEED Lipids database

Property	Model Equation	Applicable compound
Primary properties		
Normal Melting Point	3.14	All compounds in the database
Normal Boiling Point	3.14	All compounds in the database
Critical Temperature	3.14	All compounds in the database
Critical Pressure	3.14	All compounds in the database
Critical Volume	3.14	All compounds in the database
Compressibility Factor	3.14	All compounds in the database
Standard Gibbs Free Energy of formation	3.14	All compounds in the database
Standard Enthalpy of Formation	3.14	All compounds in the database
Standard Enthalpy of Fusion	3.14	All compounds in the database
Temperature dependent property		
Vapor Pressure	3.36	2 3 4 5 6 7 8 16
	Modified SRK EOS	1 9 10 11 12 13 14 15 17 18
Liquid Heat Capacity	3.42	2 3 4 5 6 7 8 16
	3.41	1 9 10 11 12 13 14 15 17 18
Liquid Viscosity	3.5	2 3 4 5 6 7 8 16 (without aromatic rings)
	3.49	1 9 10 11 12 13 14 15 17 18
Liquid Thermal Conductivity	3.35	All compounds in the database
Ideal Gas Heat Capacity	3.40	All compounds in the database
Surface Tension	3.43	2 3 4 5 6 7 8 16
	3.44	1 9 10 11 12 13 14 15 17 18
Enthalpy of Vaporization	3.38	2 3 4 5 6 7 8 16
	3.37	1 9 10 11 12 13 14 15 17 18
Liquid Density	PCSAFT (Gross and Sadowski, 2001)	All compounds in the database
Vapor Viscosity	3.46	All compounds in the database
Vapor Thermal Conductivity	3.47	All compounds in the database

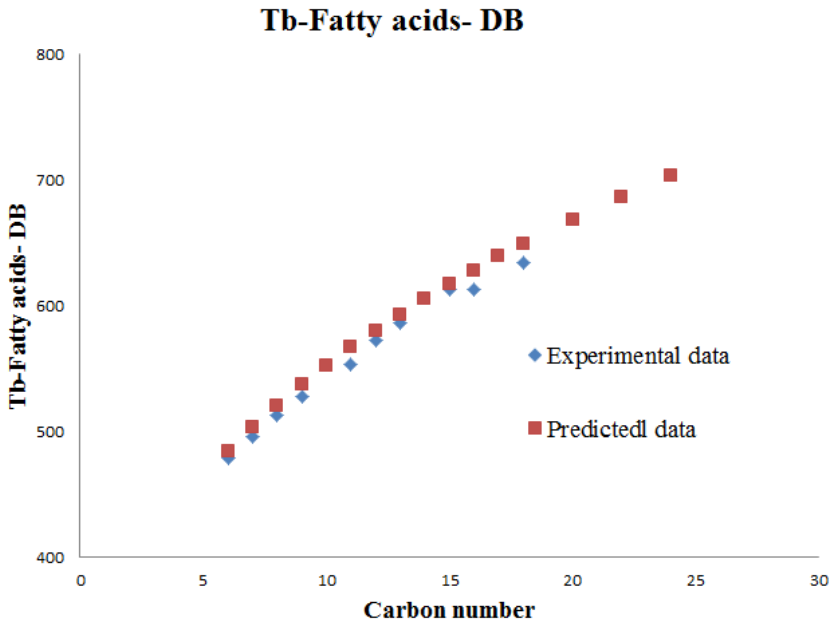


Figure 3.6. Consistency check of melting point of fatty acids

3.3.3.3 Property model creation

Eqs. (3.45) is able to calculate dynamic liquid viscosity for compounds that are commonly found in biodiesel and related oleochemicals. However, for more complex esters that can be used to produce lubricants such as dipropyl adipate and dibutil adipate (**see Figure 3.7**), their complex molecular structures were not taken into account in the model of Ceriani et al (2011). Therefore, in this work, the Ceriani et al (2011) viscosity model is extended through the following development steps:

- Collect the experimental data
- Create additional new molecular structure groups: The new groups are identified by applying 'molecular structural similarity criteria' based approach. In this approach, the molecular structure of one compound is compared with the structures of other compounds in the data-set to identify a set of compounds that are "similar" in nature. For example, didecyl adipate is similar to dibutil adipate in that, they both have two (COO) groups (see **Figure 3.7**).
- Perform data-consistency check and identify data points with large uncertainties: η is dependent on temperature. Since the measured value varies with the purity of the sample of pure compound used in the experimental measurement, the

value of η also varies significantly. This can be observed from **Figure 3.8** and **Figure 3.9** that the measured values of η of a particular class of pure compounds fall perfectly on the line when plotted against the increasing carbon number. Hence, the data-points that are not falling on the straight line may not be necessarily inconsistent data-points;

- Property function selection: this property function is selected based on the model from Ceriani et al (2011) in order to maintain the ability to estimate η of common lipids and extend the application range of the model to estimate η of complex esters;
- Parameter regression: the parameter regression of the contributions, A1k, B1k, C1k, A2k, B2k and C2k was carried out in one step. The objective function is to minimize the error defined by **Eqs. (3.93)**. The estimated values of η are well fitted with the experimental data (see **Figure 3.10**). The statistical analysis of the standard deviation (SD) (see **Eqs. (3.94)**), the relative deviation (RD) (see **Eqs. (3.95)**), the average absolute error (AAE) (see **Eqs. (3.96)**) and the average relative error (ARE) (see **Eqs. (3.97)**) are done after the regression. The results are shown in **Table 3.31**;
- Prove the capability of the model: the model is tested with a set of extra data points (2 complex ester mixtures). The results showed that the model is able to predict η of pure compounds and give high accuracy results for the prediction of η of mixtures.

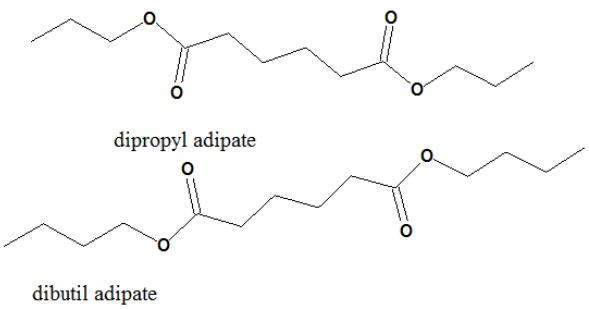


Figure 3.7. Molecular structure of dipropyl adipate and dibutil adipate

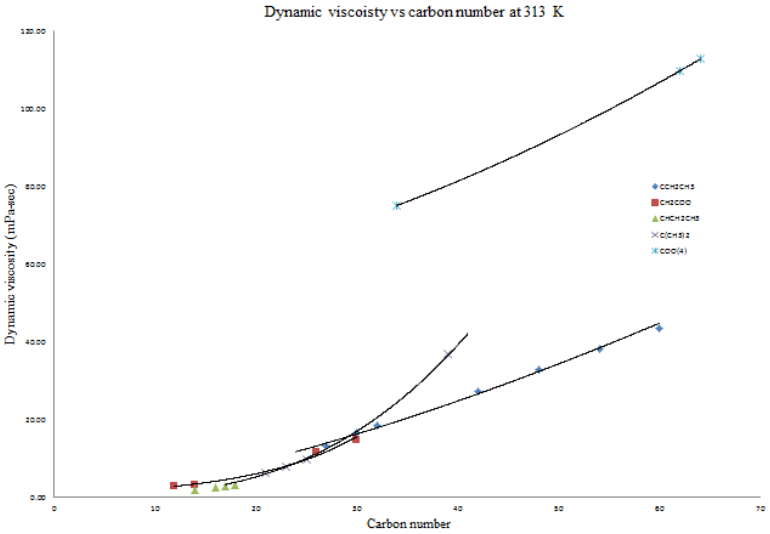


Figure 3.8. Plot of dynamic liquid viscosity versus carbon number of complex esters at 313 K

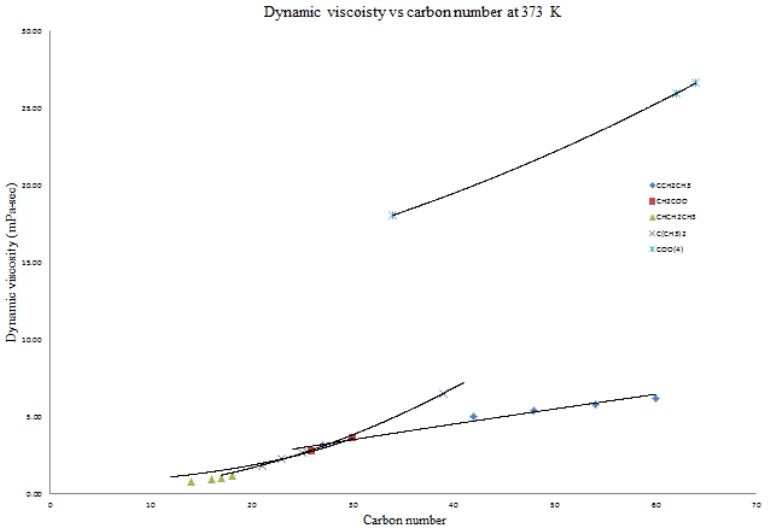


Figure 3.9. Plot of dynamic liquid viscosity versus carbon number of complex esters at 373 K

$$Objective\ function = Min\ (\eta^{Pred} - \eta^{exp})^2 \qquad (3.93)$$

$$SD = \sqrt{\frac{\sum(\eta^{Pred} - \eta^{exp})^2}{N}} \tag{3.94}$$

$$RD = \frac{|\eta^{Pred} - \eta^{exp}|}{\eta^{exp}} \cdot 100 \tag{3.95}$$

$$AAE = \frac{\sum|\eta^{Pred} - \eta^{exp}|}{N} \tag{3.96}$$

$$ARE = \frac{\sum RD}{N} \tag{3.97}$$

Table 3.26. Statistical results from the regression of η model

Statistic analysis	Value
R^2	0.9992
Average absolute error, AAE	0.50
Average relative error, ARE	3.87
Standard deviation, SD	1.05

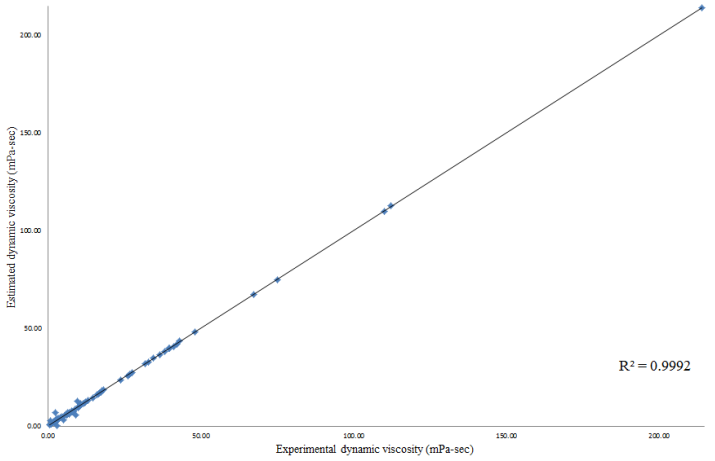


Figure 3.10. Estimated versus experimental data of dynamic viscosity

3.4 Design methodologies

3.4.1 Molecular design algorithms

3.4.1.1 Generate and test approach

All the steps of the generate-predict-select procedure are performed sequentially. Single molecular candidates are generated from a set of molecular building blocks based on UNFAC groups. The generation of molecular candidates is explained in section 3.1.1.2 in chapter 3. The properties of molecular candidates then are calculated and subsequently tested against the design specifications. This algorithm is used to avoid the combinatorial explosion, which occurs when the size of the problem becomes so large that computational time becomes too excessive.

3.4.1.2 Mathematical programming

All the steps of the generate-predict-select procedure are performed simultaneously. The molecular design problem is formulated as an optimization problem where the constraints are treated as mathematical equalities and/or inequalities and the performance indices are combined into an objective function, which is minimized through an appropriate numerical method. CAMD problems are usually formulated and solved as the Mixed Integer Non Linear Programming (MINLP) solutions (Duvedi and Achenie, 1996; Camarda and Maranas, 1999) because some property models such as flash point and liquid viscosity are non-linear models. The MINLP problem can be expressed as follows (Karunanithi et al., 2005; Zhang et al., 2015):

$$\min or \max f_{obj}(\boldsymbol{X}, \boldsymbol{N}) \tag{3.98}$$

$$structural\ constraints: g_1(\boldsymbol{N}, \boldsymbol{Y}) \leq 0 \tag{3.99}$$

$$property\ constraints: g_2(\boldsymbol{N}) \leq 0 \tag{3.100}$$

$$process\ model\ and\ other\ constraints: g_3(\boldsymbol{X}, \boldsymbol{N}) = 0 \tag{3.101}$$

$$\boldsymbol{X} \epsilon\, \boldsymbol{R^n}, \boldsymbol{N} \epsilon\, \boldsymbol{Z_+^n}, \boldsymbol{Y} \epsilon\, \{\boldsymbol{0}, \boldsymbol{1}\}^{\boldsymbol{q}}$$

$\boldsymbol{N}$ is a vector of integer variables, which are related to the numbers of the building blocks. $\boldsymbol{Y}$ is adjacency matrix which is related to the description of the molecular structure. $\boldsymbol{X}$ is a vector of continuous variables, which are related to the mixture and/or process variables.

The first constraint deals with the structural stability of the generated molecules. Here, molecules that do not follow rules such as the octet rule (valency) will be rejected as infeasible molecules (Churi and Achenie, 1997). The second constraint includes pure component property constraints. These are, for the case of group contribution methods, often linear equations. In the third property constraint the mixture properties are investigated. Here, both mixture composition dependent properties, such as mixture stability, and mixture independent properties, such as solvent selectivity, are considered. There are often nonlinear equations. In the last and fourth constraint type, process models are included, such as mole balance for a unit operation. Other models may also be included here such as phase equilibrium equations. **Eqs. (3.98)** is the objective

function, which can be: product cost minimization or product property maximization. $g_1(\boldsymbol{N},\boldsymbol{Y})$ represents the linear constraints, which are structural constraints such as maximum number of UNIFAC groups presented in a molecule candidate. $g_2(\boldsymbol{N})$ represents target properties constraints (linear or non-linear properties). $g_3(\boldsymbol{X},\boldsymbol{N})$ depends on the specific problem.

3.4.1.3 Decomposition methods

MINLP problems can be very large and difficult to be solved. Therefore, Achenie and Karunanithi (2005) decomposed the problem into four sub-problems:

- Sub-problem 1: feasible molecular structures are generated based on the structural constraints. This sub-problem is a function of the binary variables. This way, the candidates that are equal to one are selected, and those that are equal to zero are rejected;
- Sub-problem 2: the feasible molecular structures from sub-problem-1 are solved for the pure compound properties. Those molecules, which satisfy the pure component property constraints, are then passed into sub-problem 3. This sub-problem is also a function of binary variables alone (because these constraints handle only primary structure-based properties);
- Sub-problem 3: in the case that the final molecular product is to be used with other compounds (such as design of a solvent molecule for extractive distillation of an azeotropic mixture), the mixture properties such as activity coefficient and solubility of each compound are needed to be calculated. Those satisfying the mixture property constraints are passed on to sub-problem 4;
- Sub-problem 4: the process model constraints (function of both integer and continuous variables) are considered along with the objective function, giving optimal solutions by solving a smaller problem.

3.4.1.4 Database search

There are two algorithms employed as following:

- Forward approach: this approach is used when the interested chemical structure is known and chemical information (such as experimental data of properties, solubility data and solvent data). It is usually employed in the estimation of properties such as boiling point and critical temperature of n-hexane can be used to calculate liquid thermal conductivity at a certain temperature;
- Backward approach: this approach is used when the product target properties are known and the chemicals that satisfy the targets are identified. For example, a search for which compounds form azeotrope with n-hexane, a search for solvents that are immiscible with a product, a

search for toxic compounds that should not be included in the product design decision and many more.

3.4.2 Mixture/blend design algorithms

In the case that it is not possible to find a single molecule that satisfies all target properties, the mixture/blend design algorithms are employed. These algorithms need databases of feasible molecules that can be generated though CAMD. In general, a mixture/blend design problem is formulated as a mathematical programming problem as explained as follows.

3.4.2.1 Mathematical programming

In general, a mixture/blend design problem is formulated as the MINLP problem. The product performance index is optimized subject to product target properties, process specifications and/or cost. The design objective is limited by the mixture constraints, product property constraints and process model constraints. By considering multiple types of constraints, a general mixture/blend design problem is formulated as follows:

$$\min or \max f_{obj}(\boldsymbol{X}, \boldsymbol{N}) \tag{3.102}$$

$$pure\ property\ constraints: g_1(\boldsymbol{N}) \leq 0 \tag{3.103}$$

$$mixture\ property\ constraints: g_2(\boldsymbol{X}, \boldsymbol{N}) \leq 0 \tag{3.104}$$

$$process\ model\ constraints: g_3(\boldsymbol{X}, \boldsymbol{N}) = 0 \tag{3.105}$$

where $\boldsymbol{N}$ is a vector of integer variables, which are related to the numbers of the type of mixtures; $\boldsymbol{X}$ is a vector of continuous variables, which are related to the mixture and/or process variables; f_{obj} is the objective function to minimize/maximize one or more of the following parameters: the blend compositions, the type of mixture, cost, environmental impact, safety factor or product performances; $g_1(\boldsymbol{N})$ represent pure property constraints; $g_2(\boldsymbol{X}, \boldsymbol{N})$ represent mixture property constraints; and $g_3(\boldsymbol{X}, \boldsymbol{N})$ represents process model constraints.

3.4.2.2 Decomposition methods

The above mixture/blend design problem involves a large database of chemicals and non-linear constraints, which is possible to create a combinatorial explosion within a very large search space. A systematic decomposition based solution approach developed by Yunus et al., 2014 is used to efficiently manage the complexities and to reduce the search space. Main ingredients (MIs) are liquid mixtures such as gasoline mixtures and jet-fuels, which compose of various single molecules mixed together. The chemicals (additives) that are used for blending with MIs in order to make the blends satisfy the target properties. The decomposition based solution approach divides the MINLP problem into sub-problems that are relatively simple and easy to solve as follows:

- Sub-problem 1: the pure component properties of additives in the database and MIs such a gasoline mixture are compared with respect to the target values. Note that, this step is applied only for the linear target properties. **Figure 3.11** illustrates the comparison of the binary mixture, where ζ_i^k represents the target property of additive *i*; ζ_{MI}^k is the target property of *MI*; ζ_{LB}^k is the lower bound of the target property, *k*; and ζ_{UB}^k is the upper bound of the target property, *k*.

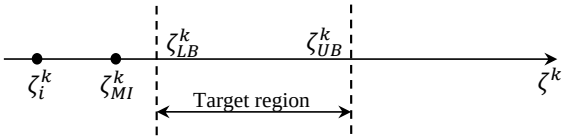


Figure 3.11. Representation of the property comparison. Binary mixture of *MI* and chemical *i* is infeasible.

For example, a binary mixture that has the pure compound property of MI and the additive *i* are both either lower than the lower bound values, or greater than the upper bound values is rejected. All the feasible additives are passed to sub-problem 2;

- Sub-problem 2: this sub-problem calculates the Gibbs energy of mixing of the mixtures at the temperature at which the miscibility test has to be performed (Conte et al, 2011). If any of the mixtures is found to be unstable, the multi-component mixture is regarded as unstable which is rejected to avoid any phase split of the blends. All the feasible mixtures are passed to sub-problem 3;
- Sub-problem 3: this sub-problem calculates the composition range for each linear target property for all mixtures that satisfy the corresponding property target values and identify the overall composition range for each binary or ternary mixture by comparing the composition ranges of all target properties.
- Sub-problem 4: this sub-problem calculate the non-linear mixture properties and for the remaining mixtures at the overall composition range and find new composition ranges that satisfy the non-linear constraints. Then, recalculate the linear target properties using new composition ranges. At this point, all the mixtures that satisfy the linear and non-linear property constraints have been identified.

3.4.3 Formulated product design algorithms

A number of compounds presented in formulated products are quite high (5-20), and the ingredients are quite different between each other. In chemical formulations, solvents, polymers, pigments, surfactants, aroma compounds, and so on are blended together. In addition, these products also have complex structures (polymers, pharmaceutical ingredients, etc) and forms (emulsion, suspensions, etc). Therefore, the design of each ingredient needs different type of algorithms. For example, in order to design an insect

repellent, the active ingredients which have the function to protect insects from human skin and additives which are able to enhance the product functions are obtained through the use of the database search, while the solvent mixture is designed through a mathematical programming.

These methodologies are based on the "define target- match target" paradigm. They employ the reverse design technique. That is, the product needs (defined by consumers or companies) are known and they are converted into a set of target properties, this set of properties are the constraints to be used to determine a set of promising candidates generated base on a specific algorithm of each product design methodology.

3.4.4 Device design algorithms

For the devices or functional products, a procedure for conceptual design is proposed by Seider et al (2009):

- Specify the product performance;
- Identify the key ingredients as well as product structure (configurations that can accomplish the target performance);
- Identify the physicochemical phenomena such as reactions and separation;
- Use the models, experiments, and available data/knowledge to identify the product specifications for meeting the desired product performance and compare product alternatives.

For example, a design of an air purifier that has a performance to remove indoor volatile organic compounds (VOCs). The key ingredients are VOCs such as toluene or acetone. The VOCs can be treated by photocatalytic oxidation (PCO). The next step is to choose the PCO catalyst, UV generator (UV lamp), a catalyst support, the structure, form, shape or configuration of the product. For example, Pt-doped TiO2 catalyst can be used as the catalyst and impregnated on high efficiency particulate air (HEPA) filter which is a highly porous mat made up of randomly oriented synthetic fibers. In order to increase the contact area, the product's shape can be a folded cylindrical filter with a UV lamp at the center instead of cubic filter. The associated phenomena is the photocatalytic oxidation of VOCs on TiO_2 surface where Langmuri-Hinshelwood kinetics can be used to model the PCO reaction in order to identify amount of TiO2 coating needs in order to reduce the VOCs concentration. The experimental test can be done to measure the concentration of VOCs of the outlet stream. Through these steps, a knowledge base and databases are needed to: identify the product specifications for meeting the desired product performance; choose key ingredients as well as product structures; and identify the associated phenomena. Therefore, Fung and Ng (2003) emphasize that the most important is to select the proper operating conditions for the processing equipment to achieve the desired product qualities. Morales-Rodriguez and Gani (2009) employ: a knowledge base containing data related to fuel-cells and microcapsule for control release design and multi-scale modeling approach to design products and to study behavior of products.

3.5 Computer-aided tools

The solution of a chemical product design problem by applying systematic integrated model-based and experiment-based methodologies require the retrieval, the calculation, the use and the management of a wide range of models and procedures, as well as very large amount of information and data. Computer-aided tools are necessary, in order to simplify the use of these procedure. In this section, different tools that have been used in this work are listed;

- Modeling tools: MoT (Model construction and analysis tool); ModDev (model development algorithms); ModTem (model template algorithms);
- Solvers: MoT; GAMs and MATLAB;
- Computer aided product design tools: ProCAMD (computer aided molecular design tool); Opt-CAMD (mathematical programming based product design); SolventPro (special tool for solvent selection-design for various solvent-based applications);
- Property prediction tools: ProPred (pure compound property prediction tool); TML (thermodynamic model parameter estimation tool);
- SustainPro: sustainable process design software (Kalakul et al., 2014);
- ECON: economic analysis software (Kalakul et al., 2014);
- LCSoft: life cycle assessment software (Kalakul et al., 2014);

MoT – Model construction and analysis tool

General idea of MoT is to introduce a modelling environment, which allows the modeler to construct the desired model and perform its analysis and application using embedded tools and methods. Model is represented as a set of mathematical equations written in a natural language, e.g. similar to the way equations are written in the books or scientific articles. Therefore, no specific knowledge of additional programming language is required (Heitzig et al., 2011). MoT information flow is shown in **Figure 3.12**. The user interface systematically guides a user through the steps of the workflow. The translated model can be solved, after satisfying mathematical consistency requirements. After the model equations are solved, it is possible to generate a COM-object of the model to transfer and use it in external software such as other modeling tools or process simulators. COM-object of other models can also be used for different terms of a model.

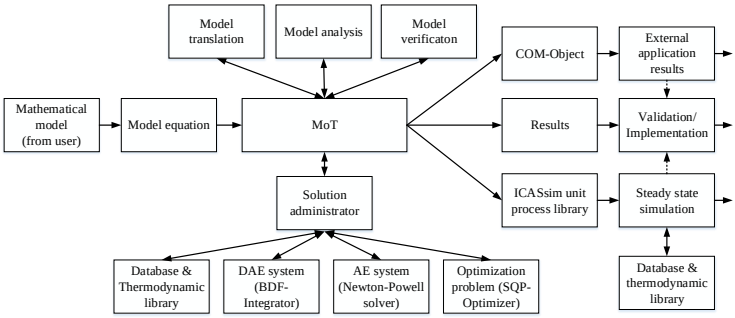


Figure 3.12. MoT information flow

ModDev – Model development algorithms

ModDev is a knowledge-based system that is able to generate the model equations for a system described by the modeler following a given structure for describing the modelling needs. In addition, it provides options for manual definition of new model building blocks and/or new variable types as well as a set of fundamental building blocks that may be combined to create the new model building blocks. The software tool provides features for visualization of the building blocks, i.e. the modeler draws the required building blocks and describes them in the corresponding forms (see **Figure 3.13**). The system automatically generates the mathematical equations, based on the description provided by the modeler. The resulting set of equations can be directly transferred to MoT or translated to the appropriate modelling language and exported to another tool. After creation and validation, the model can be as well added to the problem-specific model library, giving, thereby, opportunities for its use as a template in other model-based applications.

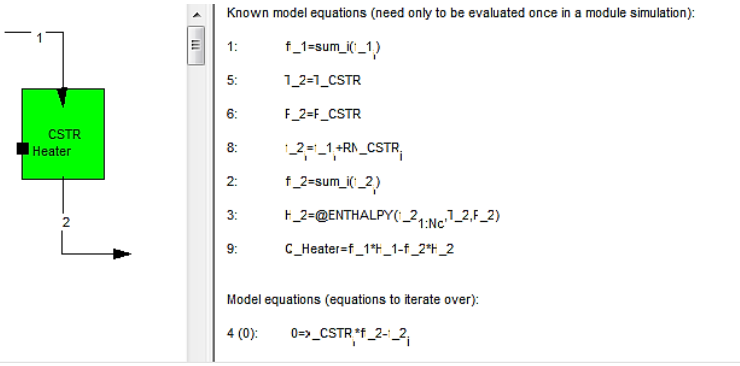


Figure 3.13. Main working area of ModDev – described system on the left-hand side and generated equations on the right-hand side

ModTem – Template tool

The template-based modelling approach is implemented as a software tool – ModTem (modelling templates), providing easy and user-friendly interface for rapid and more efficient development and use of models (Fedorova et al., 2015). It provides an environment for creation of new templates or addition of the new building blocks for existing templates and an environment for template use. The use of the software does not require additional knowledge or training, as the interface is meant to be intuitively understandable. The tool guides the user through the steps of the workflow for template-creation or template-use. It allows the user to impose changes to an existing template, to create new templates and/or prepare an existing model for use in a specific application. The final model equations are translated as an MoT-object and can be solved and identified through the MoT modelling platform or can be transferred to a text or XML-file in order to use it in external simulation environment. After model development and analysis is complete, the newly created model can be added to the model library and used as well for template creation and/or updates.

ProCAMD – Computer aided molecular design tool

ProCAMD is a tool for design single molecule products based on the CAMD technique of Harper et al. (2000). It is able to: generate the feasible molecular structures; represent the generated molecular structures for prediction of the needed properties; determine what level of molecular structural information should be used; and calculate the target properties (measured and/or predicted). In order to design a molecule using ProCAMD, five main steps with each step representing a page in the setup of the problem (see **Figure 3.14**).

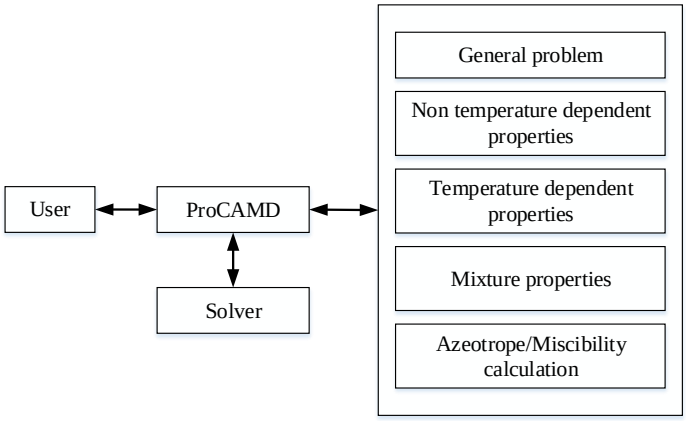


Figure 3.14. Basic steps in ProCAMD

The five options refer to:

- General problem control: this step is used to specify: class of molecules to design/search for, define the size, and complexity of the molecules (see **Figure 3.15)**;

- Non temperature dependent properties: this step is used to select non-temperature dependent properties and specify property values (see **Figure 3.16**);
- Temperature dependent properties: this option is used to specify temperature dependent properties and specify property values and temperature to be calculated (see **Figure 3.17**);
- Mixture properties: this step is used to specify mixture calculation (such as selectivity, solvent power, solvent loss, mixture viscosity, etc) wih respect to thermodynamic models that are available in ProCAMD as shown in **Figure 3.18**.
- Azeotrope/Miscibility calculations: this step is used for a azeotrope calculation (the molecule that is to be designed can be made to form azeotrope or no azeotrope) as well as a miscibility calculation (partial miscibility, total miscibility and immiscibility). The specifications of both options are shown in **Figure 3.19**.

Once the problem has been defined in terms of the property constraints, models, groups, etc., the calculation starts and the results, which is a list of molecules that satisfied all targets are identified as shown in **Figure 3.20**.

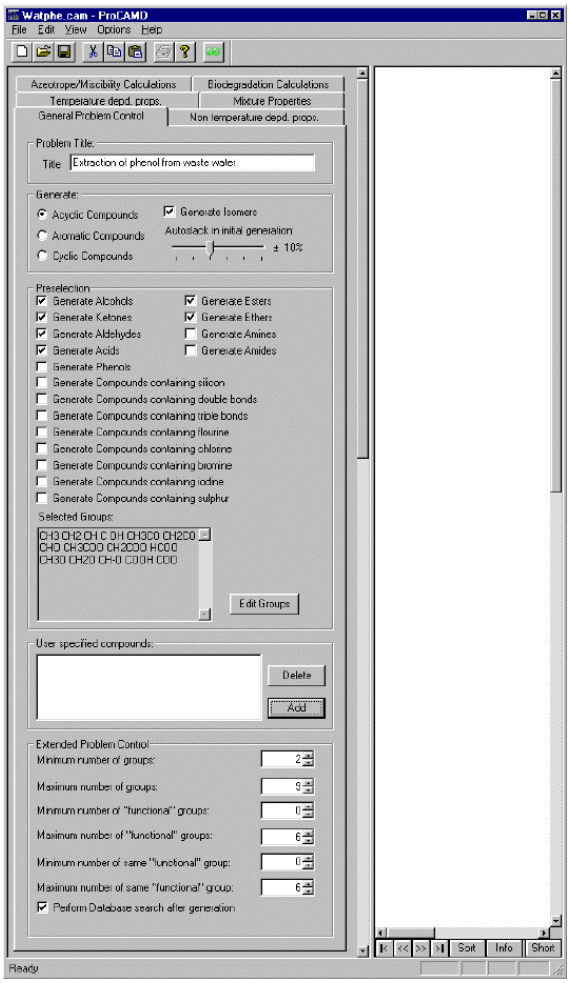


Figure 3.15. ProCAMD- General problem control page

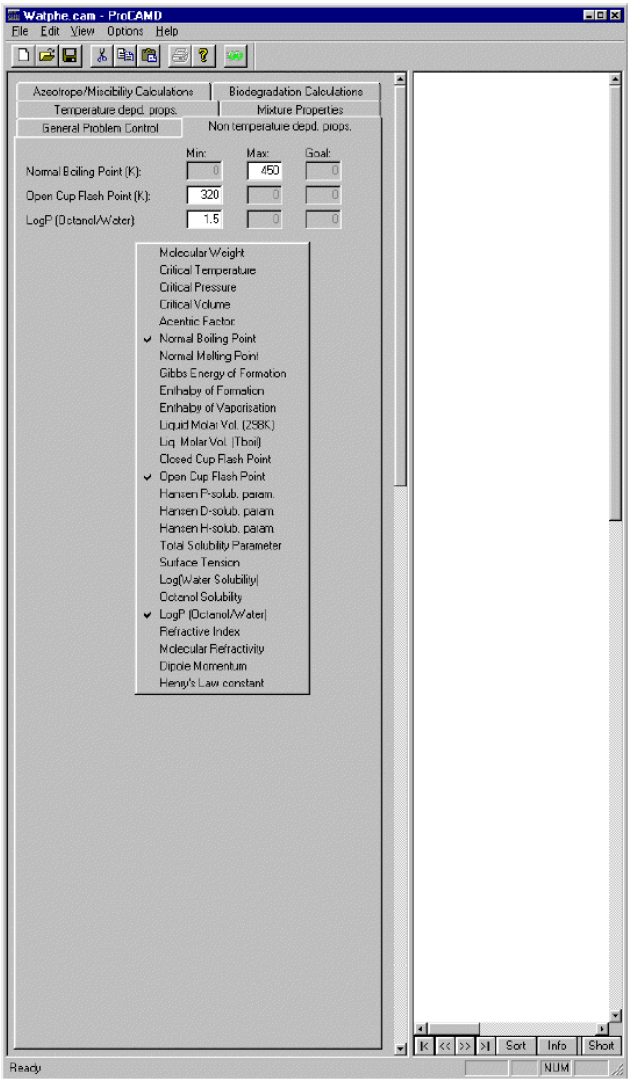


Figure 3.16. ProCAMD- Non temperature dependent property page

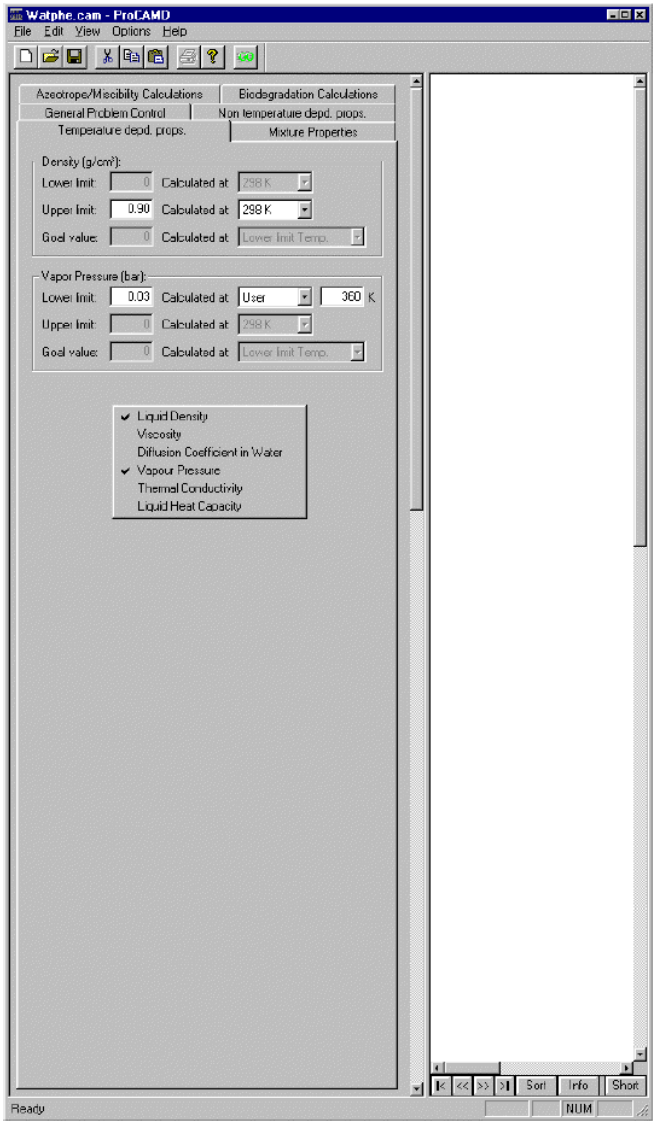


Figure 3.17. ProCAMD-Temperature dependent property control page

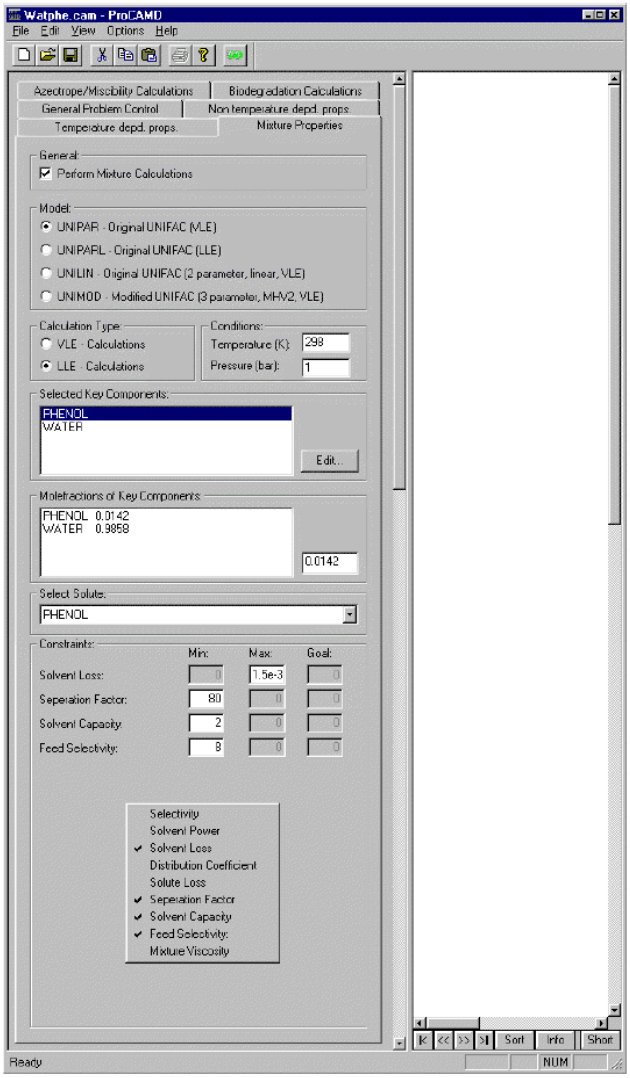


Figure 3.18. ProCAMD-Mixture property page

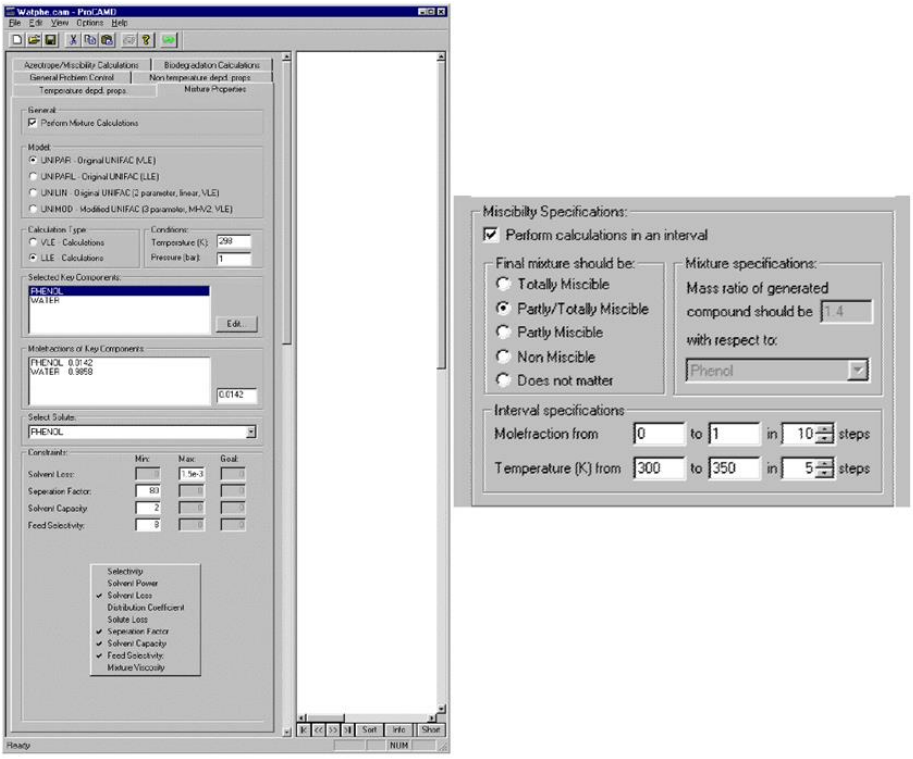


Figure 3.19. ProCAMD-Azeotrope/Miscibility calculation page

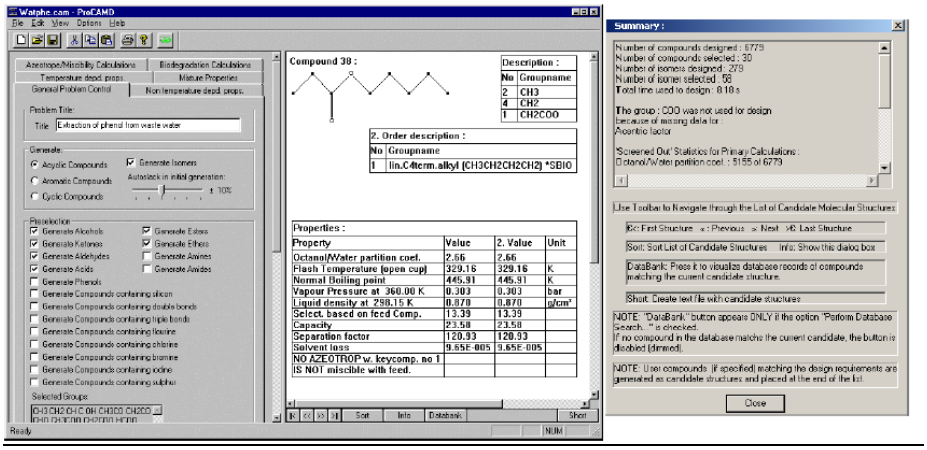


Figure 3.20. ProCAMD-Results

SolventPro – Solvent selection tool

A systematic methodology to select greener solvent for the promotion of organic reactions has already been developed and implemented in the software SolventPro as solvent selection and design template for organic synthesis. This methodology is based on thermodynamic properties of solvents, reactants and products together with the knowledge of reaction chemistry and conditions. The current module for organic synthesis is a combination of knowledge from industrial practice and computer-aided tools for property prediction and computer-aided molecular design (CAMD) principles.

ProPred – Pure compound property prediction tool

ProPred is a property estimation toolbox that allows estimation of pure compound properties based on the Marrero and Gani (2001), Joback and Reid (1987), Reid et al. (1987), and Wilson et al. (1996). The inputs for calculations of properties are chemical structure or simplified molecular-input line-entry system (SMILES) as shown in **Figure 3.21**. ProPred employes its property models to calculate primary, secondary and temperature dependent properties of the specified compounds. The interface is shown in **Figure 3.22**. It provides property values of molecules by using molecular structure as input information. For polymers, ProPred provides property predictions based on the GC+ method (combined group contribution (GC) and atom connectivity index (CI) method) and the Van Krevelen method. A database containing experimental data of pure compound properties of wide range of molecules is included in the ProPred. ProPred contains data fitting tool for fitting functional properties to several available correlations. ProPred also allows estimation of environment-related properties needed to perform calculations related to WAR algorithm and USEtox® model, and estimation of properties of lipids based on GC approach (Hukkerikar et al., 2012b).

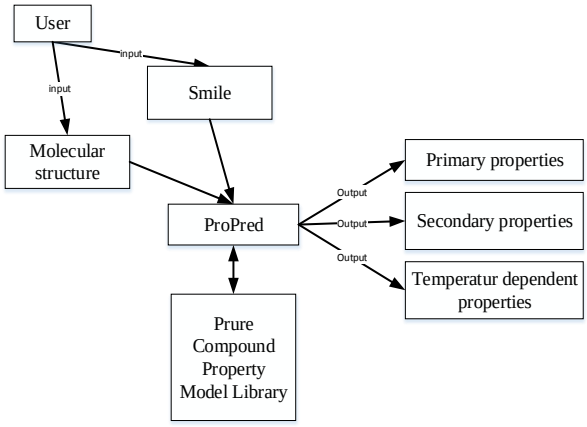


Figure 3.21. ProPred workflow/data-flow

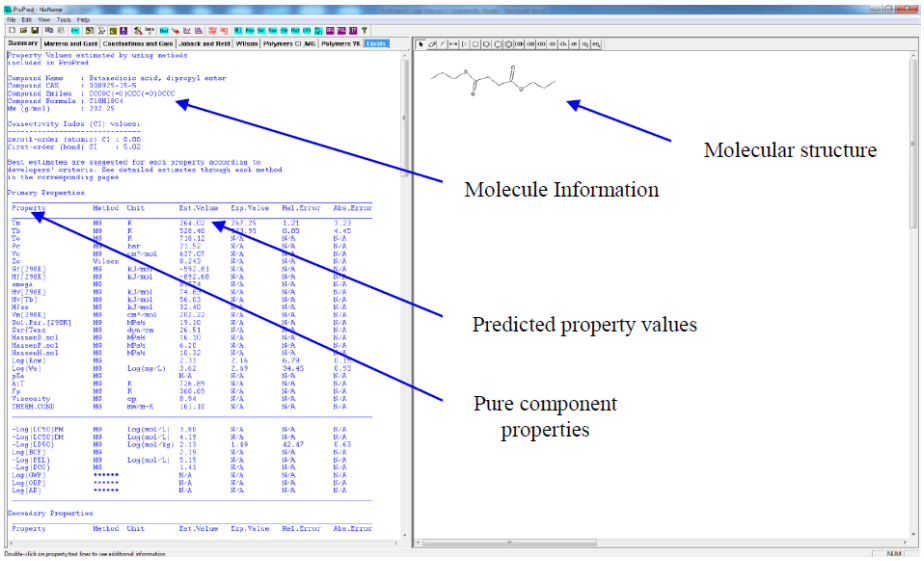


Figure 3.22. ProPred user interface

TML – Thermodynamic model parameter estimation tool

The main TML interface is shown in the **Figure 3.23**. The interface is divided into three windows. The 'Project Workspace', the 'Output Control Window' and the main document window. The 'Project Workspace' window is further divided into 'Setup', where compounds and thermodynamics can be selected, 'Regression', where thermodynamic model parameters can be determined through regression and 'Utility', where model validation and property calculations such as flash, dew/bubble points can be made. The 'Output Control Window' is used to inform the user about e.g. what to do (e.g. select compounds, select thermodynamics) and runtime information such as how many iteration a property calculation used to convergence or error messages from the thermodynamic algorithms are also shown. The main document window will at all time show the selected compounds and thermodynamics.

Opt-CAMD – Mathematical programming based product design

A generic mathematical programming formulation for computer-aided molecular design (CAMD) can be done through this tool. A given CAMD problem, based on target properties is formulated as a mixed integer linear/non-linear program (MILP/MINLP) consider considering first-order and second-order molecular groups for molecular structure representation and property estimation (Zhang, et al., 2015).

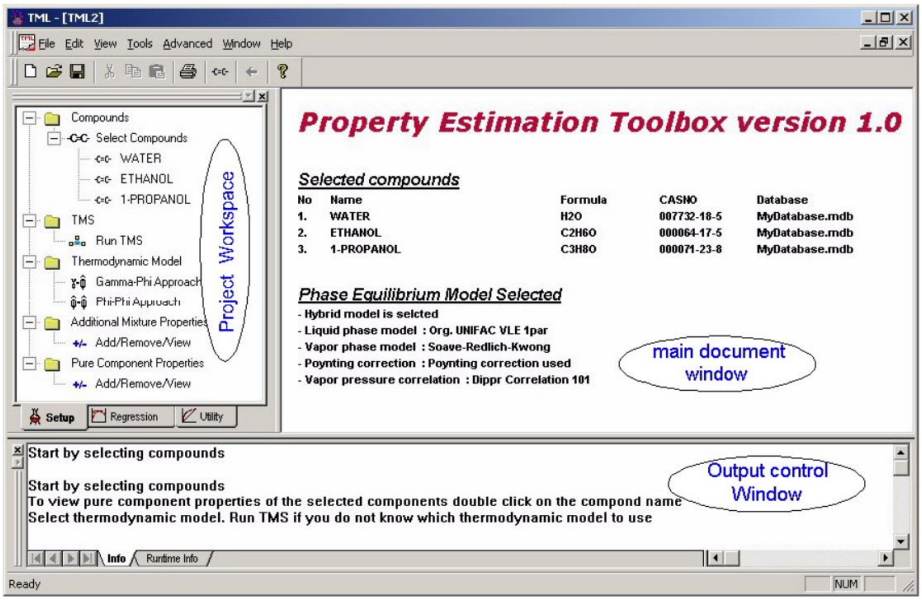


Figure 3.23. TML user interface

3.6 Computer-aided model-based framework for chemical product design and analysis

In this section, the data, methods and tools need to be organized through a framework for efficient management of the complexity. The developed systematic framework (architecture) for implementation into the product design simulator (VPPD-Lab) is shown in **Figure 3.24** where it can be seen that the framework, and therefore VPPD-Lab, handles three main product design related problems (modules). Each module is characterized by its options, algorithms, and tools, thus allowing the needed work-flow and data-flow and associated with a specific product design task to be established (see section 3.1 – section 3.5 in chapter 3).

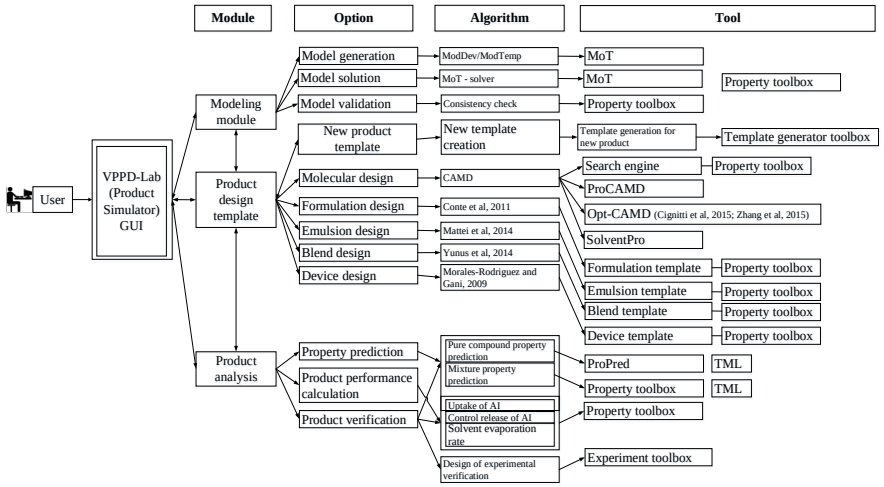


Figure 3.24. Architecture (framework) of the VPPD-Lab software

The framework allows the use of a suite of tools developed within this PhD project, such as;

- Property toolbox (see **Figure 3.25**): it is the main recurrent toolbox within the framework as different parts of it are needed by other tools (see section 3.3 in chapter 3). It consists of:

    1. Database: property data for a very wide range of chemicals found in different databases classified in terms of use in chemical products; the database also has a search engine with forward and backward search options;
    2. Property models: employs a library of property models (pure compound, mixture and phase equilibrium related properties) with associated model parameters that are linked to property prediction tools (such as, ProPred for pure compounds, TML for thermodynamic calculations, Lipids toolbox for lipids pure and mixture properties and a general property prediction toolbox covering a wide range of properties);
    3. Property prediction: employs a collection of property calculation algorithms for the needed properties (such as, phase equilibrium, solvent selectivity, liquid mixture stability);
    4. Consistency tests: checks the consistency of retrieved data from the database and/or predicted data through property models.

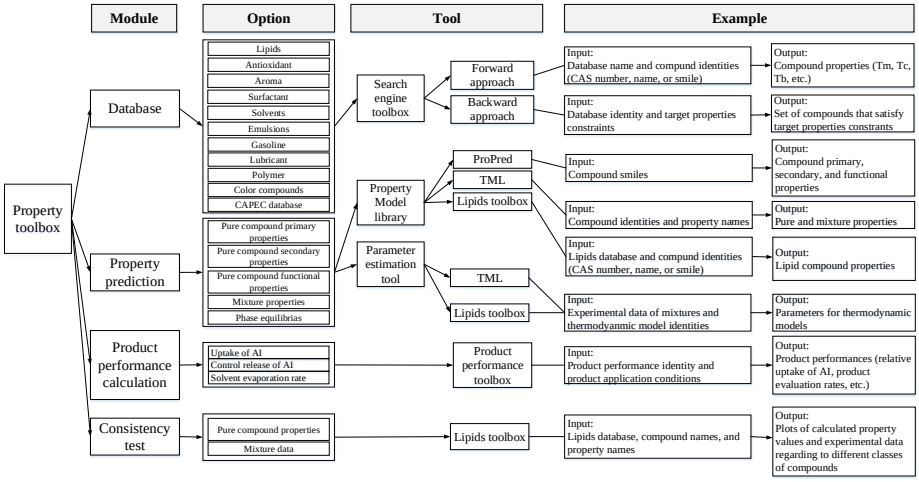


Figure 3.25. Architecture of property toolbox in VPPD-Lab

- Template generator toolbox (see section 4.2 in Chapter 4)
- Experiment toolbox for product verification and guidelines for design of experiments (see **Figure 3.3**).

Furthermore, the framework allows the link and integration with other tools, such as:

- Modeling tools: MoT (Model construction and analysis tool); ModDev (model development algorithms); ModTem (model template algorithms);
- Solvers: MoT; GAMs and MATLAB;
- Computer aided product design tools: ProCAMD (computer aided molecular design tool); Opt-CAMD (mathematical programming based product design); SolventPro (special tool for solvent selection-design for various solvent-based applications);
- Property prediction tools: ProPred (pure compound property prediction tool); TML (thermodynamic model parameter estimation tool).

4 IMPLEMENTATION OF THE FRAMEWORK

The systematic framework for chemical product design provides the user the means to solve a wide range of CPD problems in an easy and efficient manner by providing the necessary methods, tools, and references. However, it requires a number of computational tools and methods that come from different sources and disciplines. Thus, an important issue is how they can be used simultaneously and efficiently for product design-analysis? Therefore, for easy and beneficial application, the framework described in chapter 3 has been implemented into the software called VPPD-Lab (Virtual Product-process Design Software). Each part of the framework has corresponding software tools, which can communicate with each other and the data can be transferred among these tools.

The integration and merging of methods and tools from different sources have been established through the use of COM (component object model) technology and now already the .NET framework, reusable pieces of code and data in binary form that is plugged to other software components from other sources with relatively little effort. The framework implementation is able to accommodate models used for the prediction of the product property/behavior using modeling and related tools, which provides interactions with modeling engines, numerical solvers and external software. It has an interface to identify templates (work-flow) to guide users through the appropriate design-analysis steps. Special product design ontology has been developed for knowledge representation. The knowledge within each product design-analysis template is structured in terms of: product needs, their translation into properties, the corresponding property/product performance models and a wide range of data from different sources. This way, it provides a means to apply the systematic framework for chemical product design for a large range of problems. In the same way as a typical process simulator, VPPD-Lab can be routinely used to systematically, efficiently and robustly solve various types of product design-analysis problems. Available design templates, options and integrated tools corresponding to different product design-analysis scenarios are shown in the VPPD-Lab main user interface (see **Figure 4.1**). It can be noted that the same modules of the framework highlighted in **Figure 3.24**, are available in this software.

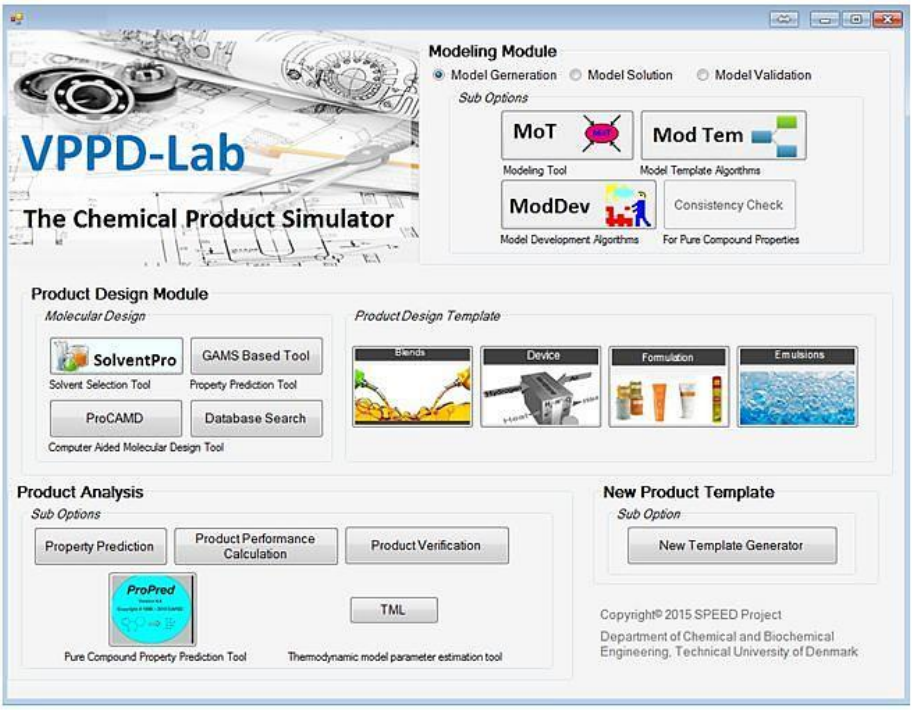


Figure 4.1. VPPD-Lab main user interface

4.1 Modeling Module

This module is designed for generation, analysis and validation of new models for product design and evaluation tasks that cannot be handled with those currently available in the models library. Examples of each option are shown in **Figure 4.2**. Through the use of computer-aided tools (MoT, ModDev and/or ModTem), it is possible to quickly create, validate and add to the new models to the model library. MoT is used when the model equations are known and the model needs to be added to the model library and also to estimate the model parameters through data regression; ModDev is used to create the new model and save it as a MoT file; ModTem is used for model-reuse, that is, take an existing model and modify it to match the objectives of the new model.

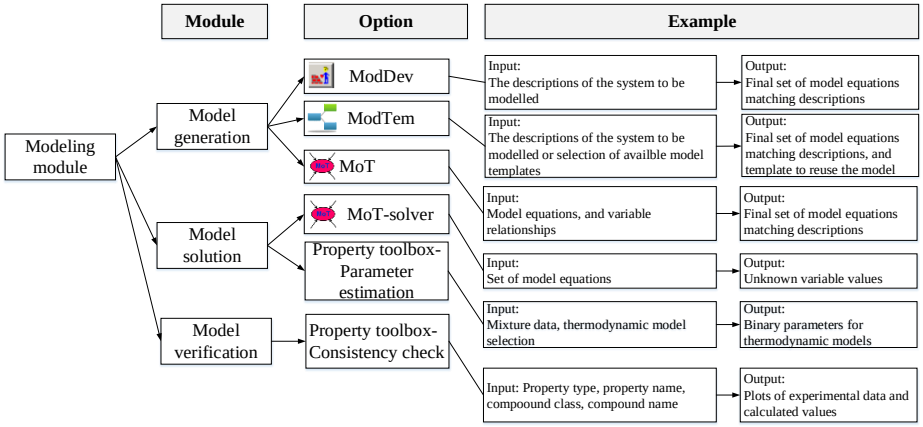


Figure 4.2. Modeling module

This module is linked to the VPPD-Lab property model library that stores property models for pure compound properties, mixture properties, and product performance as illustrated below:

- Pure and mixture property model library: Available property models are stored in an ontology (see **Figure 4.3**). The property models are divided into four main types: primary, secondary, functional, and mixture property models. Each type stores property name, information, experimental data, calculation models, model parameters for compounds that are in the VPPD-Lab database as well as new compound that is not in the database. For example, the user wants to get the higher heating value (HHV) of n-decane. The experimental data for n-decane is available in the database. Therefore, the user is able to choose the HHV values from both experimental and calculated values. For the calculated value, the user is able to choose the calculation models that are in the library (such as group contribution methods for prediction of HHV). The model parameters (such as critical properties and group parameters which are the input of the HHV model are sent to the property toolbox to calculate the HHV value. Therefore, the user is able to get information about each property model in the model library and is able to select suitable property models with respect to classes of compounds that the user is interested in (see **Figure 4.4**).

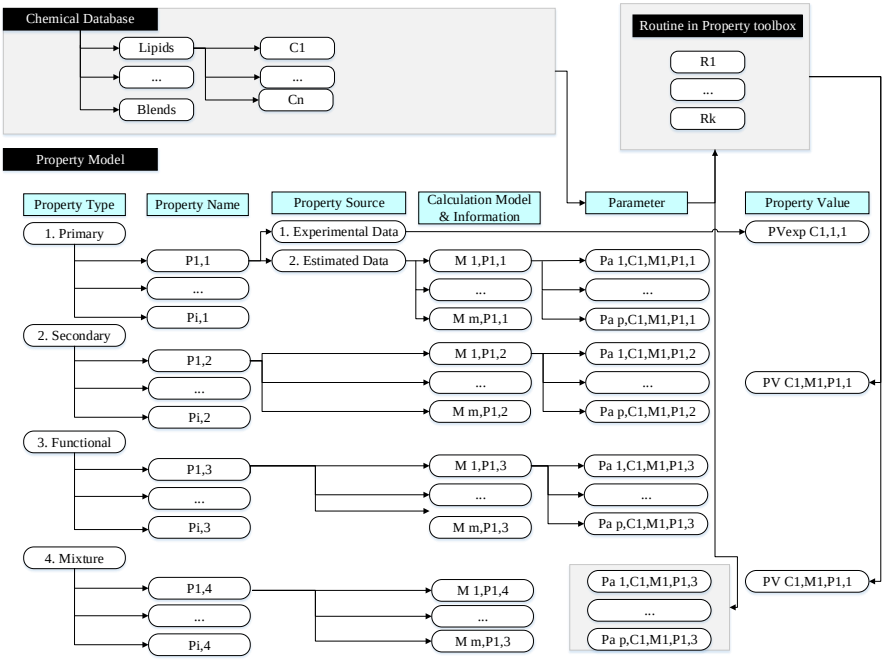


Figure 4.3. Property model ontology

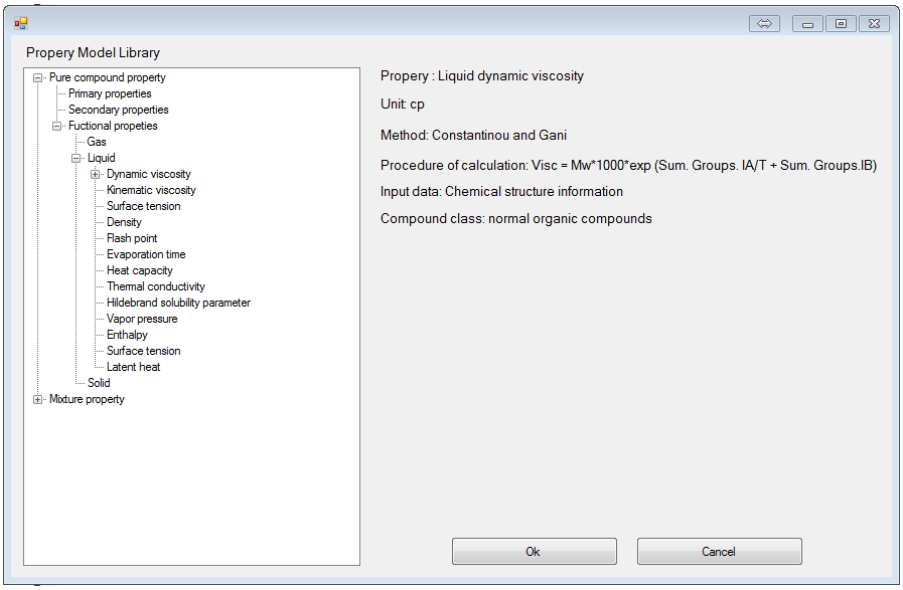


Figure 4.4. VPPD-Lab property model library for pure and mixture properties

- Product analysis (product performance) property model library (see **Figure 4.5**): The library lists available product analysis models with respect to types of product behavior. The user is able to choose to simulate product behaviors according to the ontology route. After the model is chosen, the interface guides the user to select relevant compounds and property models. Once, the compounds and the property models are defined, the software automatically reads VPPD-Lab chemical databases, simulate the product analysis problem and interprets it to the user interface view.

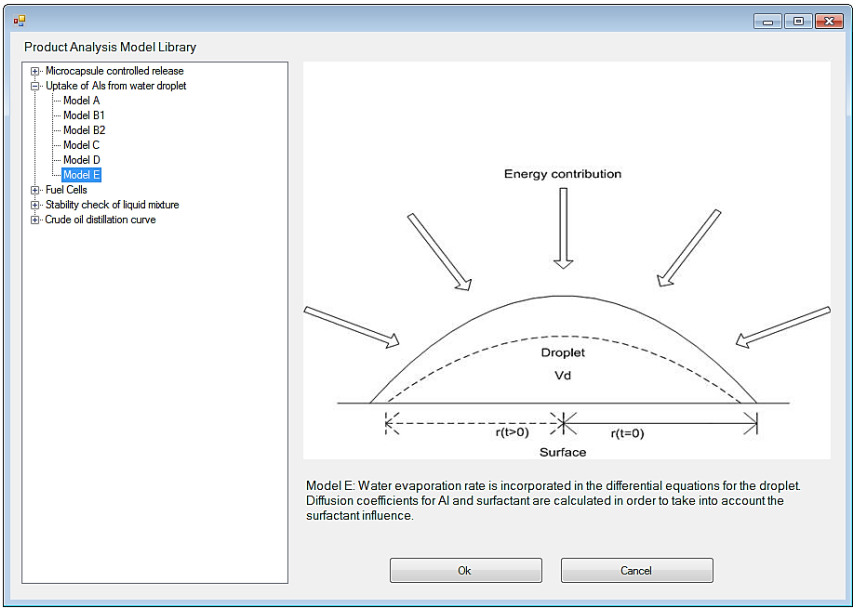


Figure 4.5. VPPD-Lab product analysis – product behavior calculation

4.2 Product design template

In this module, the template approach is used to provide the corresponding product design/evaluation workflow, the associated data-flow, tools, models, and calculation algorithms for each type of CPs.

4.2.1 Template creation

The structure of the template is shown in **Figure 4.6**. Each step in the template is incorporated with auxiliary tools. The knowledge base stores information about product design needs, target properties, molecular structural constraints, list of experiments for verification, and many more. For example, in the step 'problem definition', the knowledge base provides needs/functions for a selected type of blend (such as gasoline blends or jet-fuel blends) and translation of needs to the target properties, which will be

retrieved for calculation in the problem formulation. In the problem formulation step, the properties of the additive candidates are retrieved/and or calculated by employing property toolbox. In the problem solution step, the algorithm such as MINLP needs property toolbox to calculate mixture properties as well as a solver. Property toolbox employs databases for different classes of compounds, property models (pure and mixture), product performance models, process models that are related to a specific product design/evaluation problem, etc.

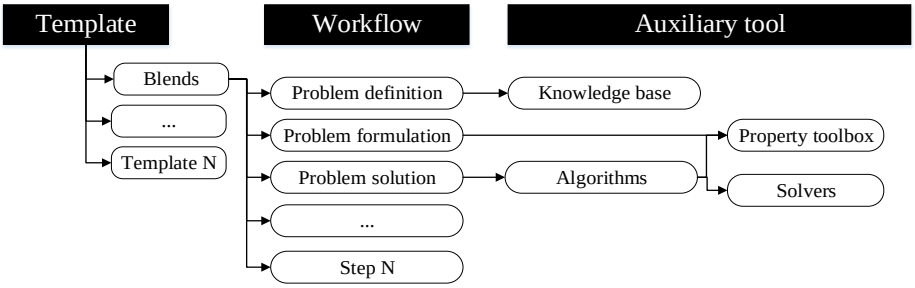


Figure 4.6. Template structure

4.2.2 Template use

After the product design templates are created and saved, they become available for utilization within the product design module. The user is able to change available options (such as chemical compounds, property models, product target property constraints) allowed by each product design template based on the product design goals. However, the user is allowed to create a new template for his special design problem. Product design and analysis template includes model equations, model description. The template generator helps the user to create user templates for a new product design or product evaluation by giving the user freedom to modify an available template to match the needs of the user specific problems or the user can follow the generic workflow for product design to create a new template, where different models and/or data may be used.

Figure 4.7 shows available product design templates to design a wide range of chemical product types, as classified by Gani and Ng (2015): (1) single molecules, (2) formulations, (3) emulsions, (4) blends and (5) devices. Each option consists of a number of problem specific templates (methodologies) together with their corresponding database, solvers, product design algorithms and analysis tools.

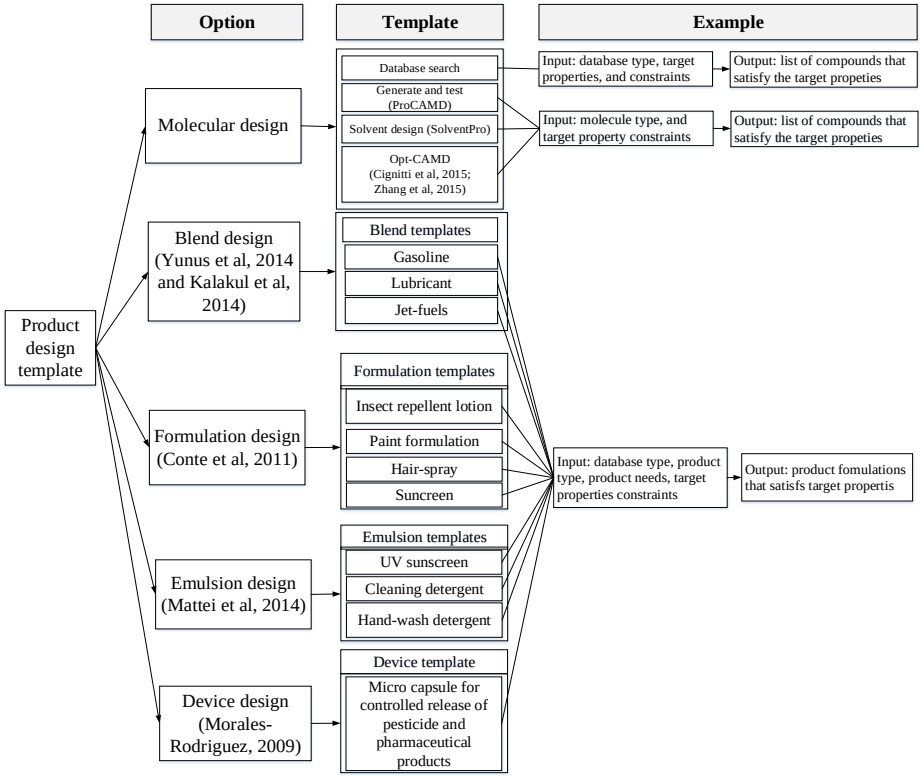


Figure 4.7. Product design module

4.2.2.1 Molecular and blend design

The design problems for these single molecule products are typically formulated by giving the specifications (properties) of the desired product (chemical) and solved by determining the molecular structures of the chemicals that satisfy the specifications, or, determining the mixtures (blends) that satisfy the desired product specifications. Blends are considered if single molecules are unable to satisfy all the desired product specifications (Gani, 2004). Common examples of single molecule products are solvents, ingredients and refrigerants where the size and complexity of the design problem depends on the size and structure of the molecular product. Different options may be used to solve these problems:

- Database search: For small molecules such as solvents and refrigerants, using the database search template can help to find the molecular structures that match the desired target properties;
- Generate and test approach: ProCAMD employs this approach (Harper et al., 2000);
- Solvent design template using SolventPro: several options that solve specific solvent selection-design-evaluation problems are available through this option.

For example, solvent-based separations, solvents for organic synthesis, solvents for pharmaceutical applications and ionic liquids selection and design as solvents (Mitrofanov et al., 2012).

- Mathematical programming approach: product design problem is formulated as a mixed integer linear and/or nonlinear programming problem and solved with an appropriate numerical solver. The template is able to generate novel pure chemical products, mixed and blended chemical products (Yunus et al., 2014). In addition, process models can be integrated for the simultaneous solution of product-process problems. Through the template, needs and target properties are defined and translated to formulate as mixed-integer non-linear programming problems that include molecular structure constraints, property constraints and process constraints. It employs optimization solvers to solve and obtain optimal product designs for the specified design problem (Cignitti et al., 2015; Zhang et al., 2015).

The design templates follow the generic workflow as shown in **Figure 4.8**.

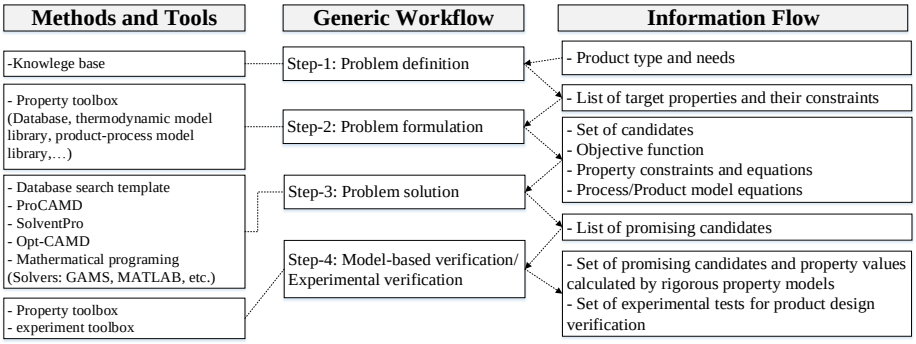


Figure 4.8. Generic workflow for blend design

Step-1: Problem definition
The product-process design problem is defined through the definition of needs and their translation to target properties for the desired products.

Step-2: Problem formulation
The computer-aided product design problem is formulated through molecular structural constraints, mixture constraints, target property constraints and process constraints. Here, knowledge base, property models, thermodynamic models and process models are also available, if needed.

Step-3: Problem solution
The design problem is formulated as a mathematical programming problem, which is usually of the MINLP type. Different solution strategies are available to solve the design problem, such as, database search, generate-test approach

(first generate candidates that satisfy the constraints and then order the feasible candidates in terms of their objective function values); simultaneous solution approach (use an appropriate optimization algorithm to solve the MINLP problem); decomposition approach (decompose the main problem into sub-problems and solve these sub-problems according to a predefined solution order). Note that it is possible to consider economic, environmental and sustainability issues through the addition of appropriate models.

Step-4: Model-based verification/Experimental verification
In this step, each feasible product design is verified either through rigorous model-based tests or experimental tests suggested by the design of experiment toolbox. In this step, the stability of the product, the desired performance, the target properties, the color, the smell, etc., are verified.

4.2.2.2 Formulation design

Formulations are the products where different chemicals are mixed to obtain a desired set of target properties. They usually contain active ingredients (AIs) that provide the main function of the product, solvents that help to dissolve and/or deliver the AIs and additives to improve the final product qualities, that is, enhance the end-use properties of the product.
Through a template implemented in the framework, it is possible to perform systematic design/verification of liquid formulated products such as insect repellent lotion, paint formulation, hair-spray and sunscreen. The template employs the formulation design methodology of Conte et al. (2011), as shown in **Figure 4.9**.

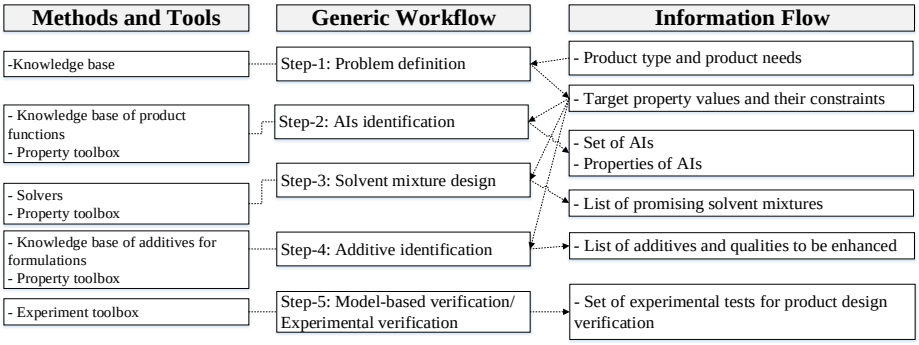


Figure 4.9. Generic workflow for formulation design

Step-1: Problem definition
The consumer needs are defined which are then translated into target properties based on a formulation knowledge base. The formulation knowledge base stores product functions (needs) and their target properties based on common sense, literature data and empirical rules suggested by experts.

Step-2: AIs identification
AIs database is searched to identify the most suitable AIs with respect to the main functions of the product.

Step-3: Solvent mixture design
Solvent mixture candidates (binary, ternary and multi-component mixtures) are generated and screened to find the feasible candidates matching the predefined target properties. In this step, a linear programing (LP) problem is first solved to identify the solvent mixture candidates subject to linear target property constraints. The feasible mixtures and their compositions then are used to calculate non-linear properties. Finally, a test for liquid phase stability of the feasible mixtures that satisfied all linear and non-linear target properties is made to identify the set of feasible solvent mixtures. Information about the stability of a liquid mixture is obtained from the calculated Gibbs energy of mixing ($\Delta G/RT$), and from its first and second derivatives.

Step-4: Additive identification
After the optimal solvent mixture is identified, the additive(s) are added in order to enhance the end-use properties of the products such as perfumes, moisturizing agents, color and neutralizing agents.

Step-5: Model-based verification/Experimental verification
In this step, each feasible product design is verified either through rigorous model-based tests or experimental tests suggested by the design of experiment toolbox. Experimental toolbox in the framework suggests a list of experimental validation tests that need to be performed to determine the physicochemical properties of the various pure compounds and/or mixtures, and to test the formulation end-use properties with respect to specific formulated products. Once the experimental validation is done, if the results do not match with the *a priori* defined target property constraints, step-2 to step-4 are repeated until all target properties are satisfied. A more practical option is to simply fine-tune (change the compositions of the additives or add a new additive) the formulation formula until all the specific constraints are satisfied.

4.2.2.3 Emulsion design

An emulsion is defined as a mixture of two or more liquids that are normally immiscible. It has suspension forms where insoluble chemicals disperse in the liquid with the help of a dispersant. Within the framework, the option for emulsions allows design of products such as sunscreen lotion and liquid hand-wash detergent where the solid constituents are emulsified through emulsifiers together with solvents and additives. The work-flow is adopted from formulation design with additional new property models, databases and knowledge base (Mattei et al., 2014). The design template follows the generic workflow as shown in **Figure 4.10**.

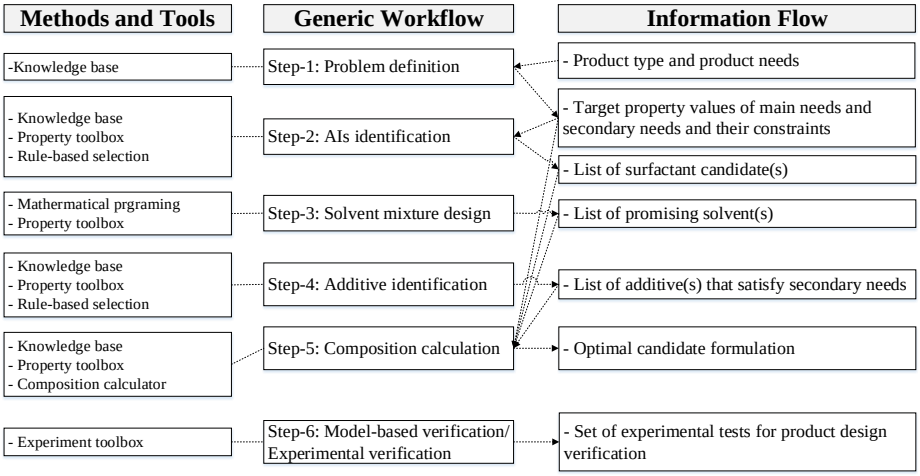


Figure 4.10. Generic workflow for emulsion design

Step-1: Problem definition
The consumer main needs and secondary needs are converted into target properties using the emulsion knowledge base. The main needs are responsible for the product main functions (such as the protection from sunburns and UV radiations, which are the main needs for a sunscreen lotion), while the other consumer needs are classified as secondary needs.

Step-2: AIs identification
Chemical candidates for AIs are screened through rule-based selection criteria, based on databases, where the relevant properties, if not available, are predicted through dedicated pure compound property models. For the emulsion design problem, surfactants and co-emulsifiers are selected as active ingredients since they act simultaneously as emulsifiers and enhance the stability of the final emulsion, respectively.

Step-3: Solvent mixture design
The design of the solvents and of the emulsifiers, driven by selection criteria based on the functional properties of the chemicals as well as consideration of effectiveness, safety, toxicity and cost, is done through a data-model based computer aided molecular design technique. Once all the ingredients have been chosen, the recipe candidates are identified through a knowledge-based mixture design method, where economic considerations are included together with appropriate boundaries related to solubility, stability, toxicity and safety issues.

Step-4: Additive identification
Additives are defined in this work as those chemicals responsible for the satisfaction of the secondary needs. A rule-base selection is applied based on the emulsion knowledge base and databases.

Step-5: Composition calculation
Once all appropriate ingredients have been chosen, the composition calculator is used to determine the overall composition of the product. In this step, the solubility of the different ingredients in the two solvents is quantified with UNIFAC-based calculations. The *a priori* defined target property constraints are considered. The knowledge base is used to set feasible composition ranges of ingredients since some of them are known to be effective only in a certain range of compositions. Finally, the emulsified product is determined by minimizing the total cost.

Step-6: Model-based verification/Experimental verification
Same as in step-5 of formulation design.

4.2.2.4 Device design

The basis of the design is to determine the device forms and constituent materials through which the desired product performances are satisfied. The challenging tasks are how to translate needs to product material properties and how constituent materials are configured since engineering science knowledge for many devices is not available. Therefore, using the knowledge base is very useful to store the information of each type of devices. The information can be product key ingredients, product structures, ingredient's property models and physicochemical phenomena models with respect to desired product performances. Device template employs a combination of different computational tools integrated within the framework: modeling tools, molecular and mixture design tool and solvent selection tool to design various types of device products. Thus, multi-scale modelling features and device design algorithms are included within the template as illustrated in **Figure 4.11** (Morales-Rodriguez and Gani, 2009).

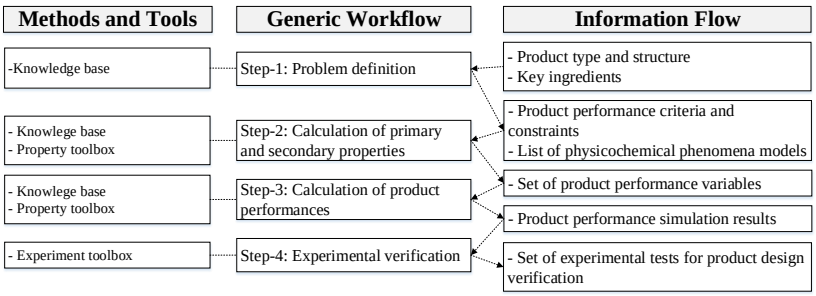


Figure 4.11. Generic workflow for device design

Step-1: Problem definition
In this step, the type of device, product performance specifications or functions, the product structure, key ingredients and main compositions as well as physicochemical phenomena models with respect to desired product performances are specified. For example, the type of a device can be

microcapsule-controlled release of pharma-products. Within the device knowledge base, the function of the product to keep an effective level of drug in the body for a specific period of time, and thereby side effects generated by drug overdosing and/or under-dosing may be avoided. The product structure (see **Figure 4.11**) contains: (1) active ingredients, which are drug molecules such as antibodies, antioxidants or probiotics that are placed in the core of the device, (2) donor medium which is a solvent that dissolve active ingredients, (3) microcapsule wall (for example, a polymer membrane) that encapsulate active ingredients and the solvent and (4) release medium which depend on the application field (for example, blood or some other medium found within the human body).

Finally, the list of product phenomena/performance models, variables and parameters for the mass transfer (by diffusion) calculation of the active ingredient through the wall are retrieved from the databases depending on classes of ingredients.

Step-2: Calculation of primary and secondary properties
Once the necessary variables are retrieved from the databases, if there are some missing primary or functional properties, property toolbox is employed to fill the gap.

Step-3: Calculation of product performances
In this step, a simulation of product performances is performed. For example, if the product is a microcapsule for controlled release, the diffusion mass transfer phenomena in the micro-scale and the microcapsule-based controlled release in the meso-scale are performed. Meso-scale calculations involve number of microcapsule sizes, number of particles, and surface area of the microcapsule to name a few. This information is used in the micro-scale to calculate the mass of released active ingredients. Within the released mass of active ingredients, the calculation of the total amount released to the receiver medium as well as the percent of the active ingredient released from the microcapsules can be calculated.

Step-4: Experimental verification
After product performances are calculated, the product is tested to verify if it satisfies the specifications. The calculation of the total amount released to the receiver medium of AI is compared with the experimental results (see **Figure 4.12**). The three scenarios use different polymer wall thickness values (scenario 1< scenario 2< scenario 3).

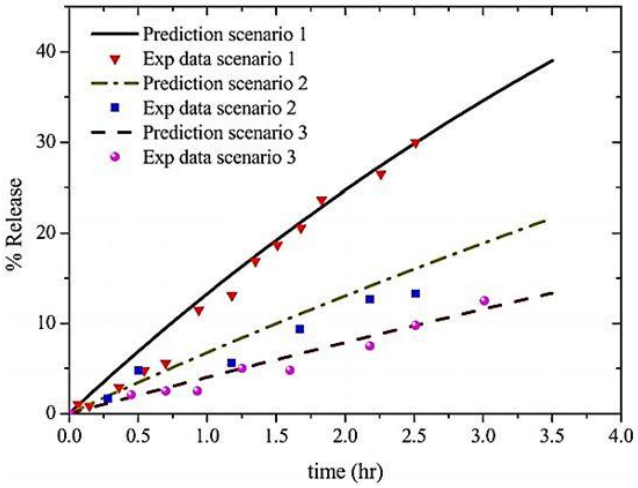


Figure 4.12. Comparison between the experimental values and models predictions of the release of codeine (active ingredient) (Morales-Rodriguez et al., 2011)

4.3 Product analysis tools

This option is used for a known product whose properties and/or performance need to be tested and/or verified. Therefore, the methods and tools used here are the same as the ones used for the model-based verification step in the generic workflow to design single molecules and blends as well as calculations in the product performance step in the device design workflow. In product analysis module, various property prediction and product behavior/performance models have been implemented into the product analysis template, such as, pure and mixture property calculations, stability check of mixtures, uptake of active ingredient, control release of active ingredient and solvent evaporation rate calculations.

Figure 4.13 shows screen shot of property prediction options in property toolbox. The prediction of pure and mixture properties of a mixture are performed. The toolbox allows to employ property models and thermodynamic models as listed in section 3.3.3 in chapter 3.

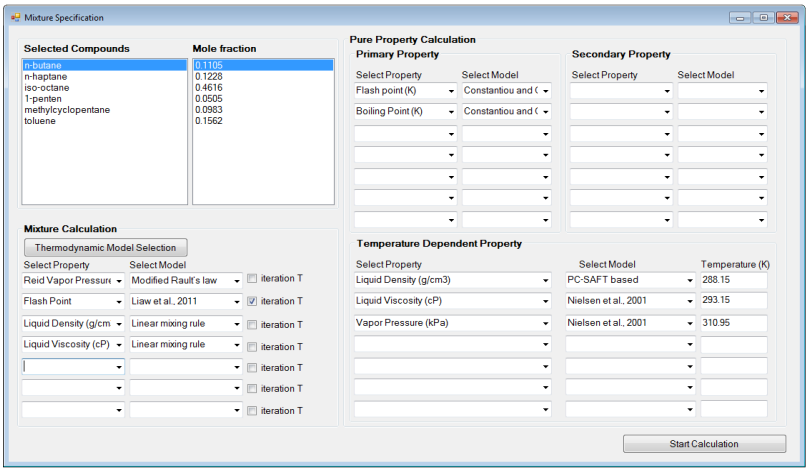


Figure 4.13. VPPD-Lab product analysis – property prediction

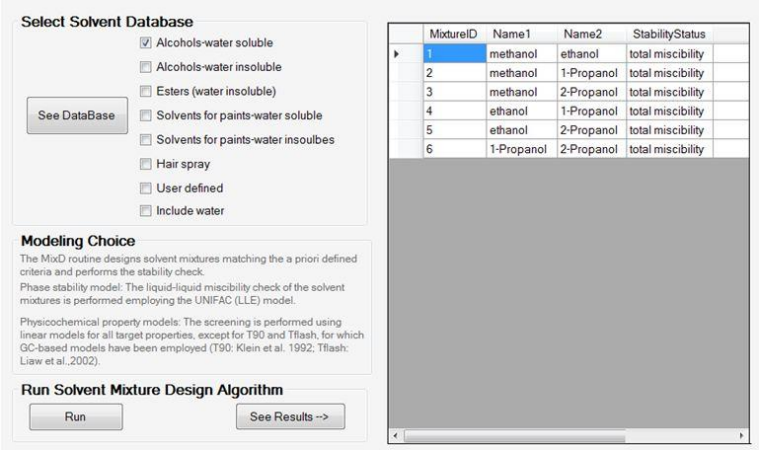


Figure 4.14. Results of stability check

Mixtures of solvents are used in many chemical-based products. The stability of these solvent mixtures is very important since these mixtures need to be stable liquid mixtures. **Figure 4.14** shows screen shot of the stability check of mixtures of binary solvent mixtures that are water-soluble and completely miscible. All binary mixtures of water plus solvent that are totally miscible are generated. The stability check algorithm is then employed to verify the solvent miscibility, giving all the possible binary combinations between the water-soluble alcohols. The solvent mixtures can be further screened in term of property such as evaporation rate, density, and many more.

4.4 XML transfer for tools integration and export/import of VPPD-Lab databases

Additional features of VPPD-Lab databases include integration of computer-aided tools within the framework and transfer chemical property data/and or models among the integrated tools as well as possibility to transfer modelling information to the external tools such as Pro/II process simulator for process design and synthesis purposes or, in opposite way to transfer data from external tools (see **Figure 4.15**). This is implemented through XML (Extensible Markup Language) transfer. The specific XML schema contains a formal description of the structure and elements in the XML files representing chemical property data. Such XML file includes single value property data within the specified units, information about property model equations and their coefficient and indexes.

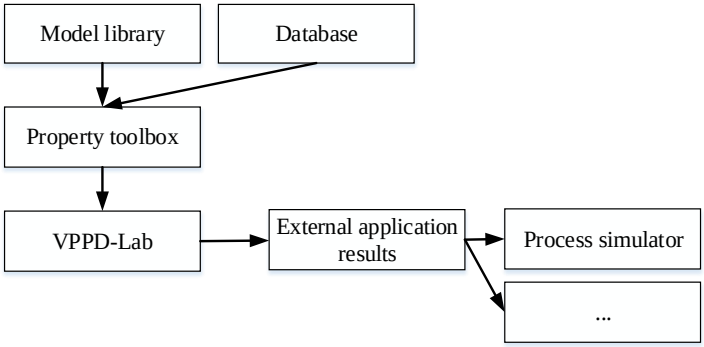


Figure 4.15. VPPD-Lab: Exporting data

An example is illustrated through the use of Lipids database in VPPD-Lab. Lipids are classified as naturally occurring molecules such as fats, waxes, sterols, fat-soluble vitamins, acylglycerides, phospholipids and many more. They are typically encountered in processes handling lipids, such as in edible oil production, production of crude gaseous or liquid fossil fuels, biodiesel production, and pharmaceutical product manufacturing. Commercial process simulators, which are increasingly used in process design and analysis, usually lack the necessary physical and thermodynamic property data and/or models for many of the lipids in their databases. Therefore, in order to design a process related to lipids, property data of lipids have been created in this PhD project and/ or collected from literatures and implemented in the VPPD-Lab property model library. The relevant property models are used to generate property data such as molecular weight, normal boiling point and critical properties as well as temperature dependent correlations such as vapor pressure, liquid density, liquid heat capacity, and other transport related properties. The generated information is stored in Lipids database and can be exported to external tools as an XML file (see **Figure 4.16**). Through the use of batch file, the commands that Pro/II accepts to create a new user chemical library "LIPIDS" containing all lipids in VPPD-Database (see **Figure 4.17**).

```
<!-- Package: LIPIDS V7.0  Date: 7 Sep 2015  -->
<!-- Added by: SK  -->
<!-- Change history: All primary and temperature dependent are checked and added -->
<comp>
<id number="21000068" name="1,2,3-tridecanoyl-sn-glycerol" formula="C33H62O6"> C-C-C </id>
<property name="MW" value="554.84"/>
<property name="NMP" value="368.28"/>
<property name="NBP" value="763.77"/>
<property name="TC" value="901.63"/>
<property name="PC" value="594"/>
<property name="VC" value="1.94406"/>
<property name="ZC" value="0.1544"/>
<property name="GFORMATION" value="-724390"/>
<property name="HFORMATION" value="-1705910"/>
<property name="HFUSIONNMP" value="91330"/>
<property name="ACENTRIC" value="1.729"/>
<property name="SOLUPARA" value="16.31"/>
<property name="FLASHPOINT" value="675.36"/>
<property name="SG60F" value="0.761396379633522"/>
<property name="MVOL25C" value="0.58866"/>
<property name="DIPOLE" value="0"/>
<property name="VDWA" value="19.452"/>
<property name="VDWV" value="23.6906"/>
<correlation name="VaporPressure" equation="20" tempUnit="k" propUnit="pa" logBasis="ln" tMin="400" tMax="670"> -149.7976 -2007.347 25.1418 </correlation>
<correlation name="LiquidEnthalpy" equation="1" tempUnit="k" propUnit="kj/kg" logBasis="ln" tMin="275" tMax="900.63"> -251179.301786813 773.117644926362 0
<correlation name="LiquidDensity" equation="1" tempUnit="k" propUnit="g/cm3" logBasis="log" tMin="368.38" tMax="900.63"> 0.994613253639601 -1.005085265273
<correlation name="LiquidThermal" equation="1" tempUnit="k" propUnit="w/m-k" logBasis="ln" tMin="368.28" tMax="763.77"> 0.216669119035185 -2.00004311572158
<correlation name="SurfaceTension" equation="1" tempUnit="k" propUnit="dyne/cm" logBasis="log" tMin="368.3768" tMax="898.5732"> 45.5794743094764 -5.560554
<correlation name="IdealEnthalpy" equation="1" tempUnit="k" propUnit="kj/kg" logBasis="ln" tMin="275" tMax="900.63"> -159703.171531309 359.447039034348 0.
<correlation name="LatentHeat" equation="15" tempUnit="k" propUnit="j/kg" logBasis="log" tMin="368.38" tMax="900.63"> 120.323427317701 -2.45044292375696 8
<correlation name="LiquidViscosity" equation="20" tempUnit="k" propUnit="cp" logBasis="ln" tMin="383" tMax="713"> -44.71737 4219.724 6.141274 -0.004133785
<correlation name="VaporViscosity" equation="1" tempUnit="k" propUnit="cp" logBasis="ln" tMin="368.78" tMax="1000"> 3.22547980122249E-04 9.026021156779078
<correlation name="VaporThermal" equation="19" tempUnit="k" propUnit="w/m-k" logBasis="ln" tMin="763.87" tMax="1000"> -0.410655445096141 0.560786719305752
</comp>
```

Figure 4.16. XML file containing chemical property data of 1,2,3-tridecanoyl-sn glycerol

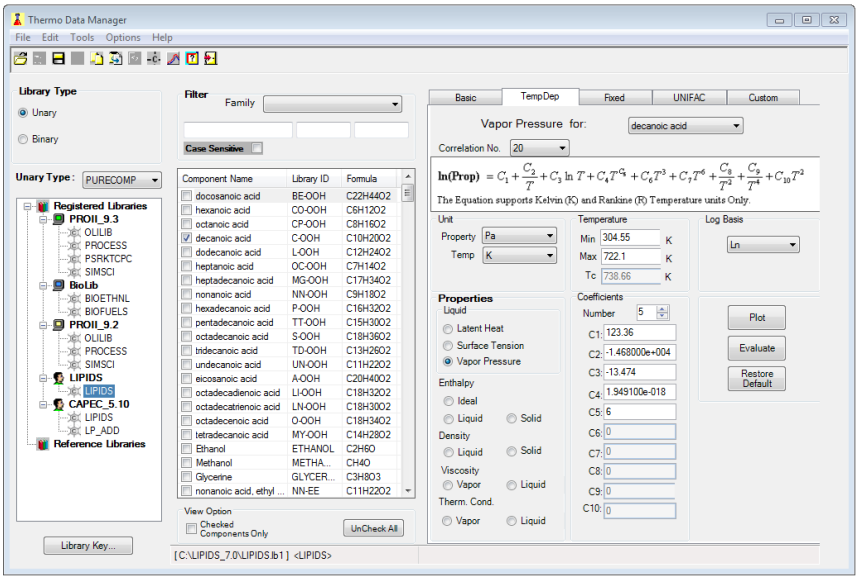


Figure 4.17. Lipids user library in Pro/II chemical library

5 APPLICATION EXAMPLES

This chapter presents case studies highlighting different features of the implemented CPD framework. The case studies are grouped into two main classes: class-1: Product analysis and class-2: chemical product design (CPD) case studies.

The application of class-1 is highlighted with:

- Prediction of mixture properties

The application of class-2 case studies is grouped in term of product classes, the design of the following products are highlighted:

- Design of molecular product
  - Solvent for extractive distillation
- Design of blends
  - Jet-fuel blends (2 case studies)
  - A diesel blend
- Design of formulation
  - An insect repellent lotion
- Design of emulsion
  - A hand-wash detergent
- Design of device

5.1 Product performance analysis

5.1.1 Verification of the properties of a gasoline mixture

A conventional gasoline comprises of a large number of hydrocarbons and for a better understanding of fuel behavior in the combustion engine, a set of properties need to be analyzed. However, a model-based surrogate mixture is used to represent it as a mixture of compounds. The objective of this case study is to analyze the properties of the surrogate gasoline proposed by Yunus et al. (2014). The composition of the surrogate is shown in **Table 5.1**.

Table 5.1. Gasoline surrogate composition

Chemical	Composition (molar fraction)
n-butane	0.1105
n-heptane	0.1228
Iso-octane	0.4616
1-pentene	0.0505
methylcyclopentane	0.0983
toluene	0.1562

The properties of interest and their corresponding models are retrieved from the property model library (see Chapter 3) of the property toolbox. ASTM distillation temperature is a set of increasing temperatures at which fuel evaporates for a fixed series of increasing volume percentages (10 percent (T10), 50 percent (T50), 90 percent (T90). It is calculated using the method of Hoffman (1969). The pure component properties are obtained from the experimental data if they are available or estimated using the models listed in **Table 5.2**. The mixture property models used are listed in Table **5.3**. The calculation results are compared with the experimental verification in **Table 5.4**. (-) means the experimental verification has not yet been done. Note that the experimental verification has been done in collaboration with TEES Gas & Fuels Research Center, Texas A&M University at Qatar, 23874 Doha, Qatar.

Table 5.2. Pure component property models

Pure component property	**Equation**
Higher heating value, HHV_i	(3.18) Yunus et al., 2014
Density, ρ_i	(3.52) and PC-SAFT (Gross and Sadowski, 2001)
Vapor pressure, P_i^{sat}	(3.53) and PC-SAFT (Gross and Sadowski, 2001)
Open cup flash point, $T_{f,i}$	(3.18) Hukkerikar et al., 2012a
Lethal concentration, $LC_{50,i}$	(3.18) Hukkerikar et al., 2012b
dynamic viscosity, η_i	(3.59)

Table 5.3. Mixture property models

Target property	Equation
Dynamic viscosity at 20 °C, η	(3.64)
Higher Heating Value, *HHV*	(3.64)
Density at 15 °C, ρ	(3.64)
Reid Vapor Pressure, *RVP*	(3.65)
Open cup flash point, T_f	(3.66)
Toxicity parameter, -log(LC_{50})	(3.64)
ASTM distillation temperature (T10, T50 and T90)	(3.73)

Table 5.4. Property calculation and experimental verification of the surrogate conventional gasoline mixture

Property	Model (VPPD-Lab)	Experimental	RSD (%)
RVP (kPa)	54.0	55.2	4.04
T_f (K)	257	-	-
ρ (g/cm^3)	0.7260	0.7113	1.45
η (cP)	0.51	0.50	1.40
HHV (MJ/kg)	45	-	-
-log(LC_{50})	3.08	-	-
T10 (K)	345	345	0.00
T50 (K)	372	372	0.00
T90 (K)	382	384	1.09

In Table 5.4, T10, T50 and T90 are the temperature of the surrogate gasoline after 10%, 50% and 90% (volume %) evaporation. The results in **Table 5.4** also show the comparison of the calculated properties with experimentally measured data. It can be noted that the model prediction accuracy is quite good. This means that these models can be employed for gasoline blend design.

Results from **Table 5.4** indicate that the gasoline surrogate can be used to represent the conventional gasoline. RVP of the surrogate is suitable for use in engines because a gasoline that has *RVP* between 45-103 kPa tends to have lower evaporative emissions (Hadfield and Dorries, 2010). Typically for gasoline fuel, the value of density at 20 ^{0}C lies between 0.720 g/cm^3 and 0.775 g/cm^3, whereas dynamic viscosity at 20^{0}C lies between 0.3 cP and 0.6 cP. However, in order to improve the surrogate, it can be blended with the higher HHV additives to increase the engine efficiency and higher T_f for the safety purpose.

5.1.2 Analysis of a surrogate jet-fuel

The objective of this case study is to analyze the properties of a surrogate jet-fuel (Agosta et al., 2004). The composition of the surrogate mixture is given in **Table 5.5**.

Table 5.5. Jet fuel surrogate composition

Chemicals	Composition (molar fraction)
dodecane	0.07
Decane	0.65
methylcyclohexane	0.01
butylbenzene	0.27

The properties of interest and their corresponding models are retrieved from the property model library (see Chapter 3) of the property toolbox. ASTM distillation temperature is the set of increasing temperatures at which fuel evaporates for a fixed series of increasing volume percentages (10 percent (T10), 50 percent (T50), 90 percent (T90). It is separately calculated using the method from Hoffman (1969). The pure component properties are obtained from the experimental data if they are available or estimated using the models given in **Table 5.6**. The mixture property models are given in Table **5.7**. The calculation results are shown in **Table 5.8.**. In order to improve the combustion performance, CO_2 emission from a combustion engine when fuels are used is estimated through a model available at USEPA (Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990 – 2005, 2007).

Table 5.6. Pure component property models

Pure component property	**Equation**
Higher heating value, HHV_i	(3.18) Yunus et al., 2014
Density, ρ_i	(3.52) and PC-SAFT (Gross and Sadowski, 2001)
Vapor pressure, P_i^{sat}	(3.53) and PC-SAFT (Gross and Sadowski, 2001)
Open cup flash point, $T_{f,i}$	(3.18) Hukkerikar et al., 2012a
Lethal concentration, $LC_{50,i}$	(3.18) Hukkerikar et al., 2012b
dynamic viscosity, η_i	(3.59) and (3.49)
Melting point, T_m	(3.18) Hukkerikar et al., 2012a
Carbondixoide emission from fuel combustion, CO_2E	(3.21)

Table 5.7. Mixture property models and their function

Target property	**Model**
Dynamic viscosity at -20 °C, η	(3.64)
Kinematic viscosity, ν	(3.64) and , $\nu = \frac{\eta}{\rho}$
Higher Heating Value, *HHV*	(3.64)
Density at 15 °C, ρ	(3.64)
Reid Vapor Pressure, *RVP*	(3.65)
Open cup flash point, T_f	(3.66)
Toxicity parameter, -log(LC_{50})	(3.64)
Distillation curve	(3.73)
Carbondixoide emission from fuel combustion, CO_2E	(3.64)
Mixture miscibility	Stability check algorithm

Table 5.8. Property calculation of the surrogate conventional jet-fuel mixture

Property	Model (VPPD-Lab)
HHV (MJ/kg)	46.48
Tf (K)	320.9
RVP (kPa)	0.53
ρ (kg/m3)	771.67
V (cst)	2.56
Tm (K)	221.15
-logLC50 (mol/L)	4.513
CO_2E (kgCO2/mile)	23.56

Results from **Table 5.8** indicate that the jet-fuel mixture can be used to represent the conventional jet-fuel. The standard requirements for worldwide jet-fuel are listed in **Table 5.10**. HHV of the surrogate is very high (typically 40 MJ/kg). T_f is higher than the benchmark (311.5 K). From the distillation curve in **Figure 5.1**, the vapor pressure of the surrogate is in the acceptable range compared to the conventional jet-fuel (Chuck and Donnelly, 2014). However, ρ should be increased to reach the aviation standards as well as CO_2E that should be reduced since globally aviation emissions contribute approximately 2& of all anthropogenic CO_2 (Chuck and Donnelly, 2014).

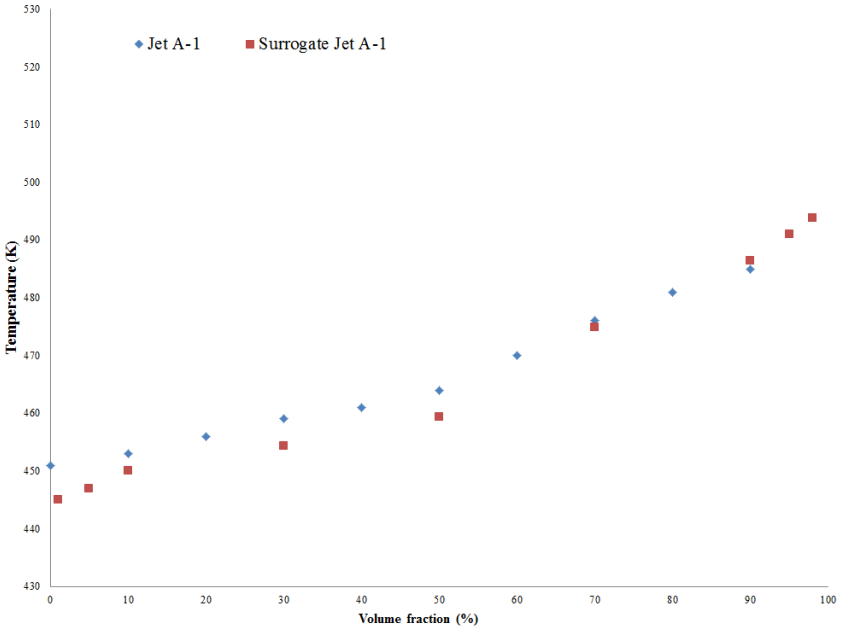


Figure 5.1. Distillation curves for the surrogate jet-fuel and the conventional jet-fuel

5.2 CPD of single species products: solvent for separation by extractive distillation

Solvent-based separations are employed for different separations. This system forms azeotrope mixture as indicated by Vapor-liquid equilibrium (VLE) calculation from the property model toolbox using UNIFAC model at 1 atm (see **Figure 5.2**). This system has to be separated into pure products. The aim of the design problem is to generate a feasible solvent list for separation of n-hexane and ethyl acetate applying extractive distillation. The solvent design problem is developed and solved using the generic workflow for molecular design as shown in **Figure 4.6**. However, the objectives of this example are to: highlight the application of VPPD-Lab knowledge base for translation of product needs into target properties (see step-1) and to demonstrate the application of different CAMD tools that help to generate solvent candidates (see step-2). The CAMD tools: "Advanced Search" option, the "Azeotropic Search" options and ProCAMD are highlighted. Note that step-3 and step-4 are not performed for this example.

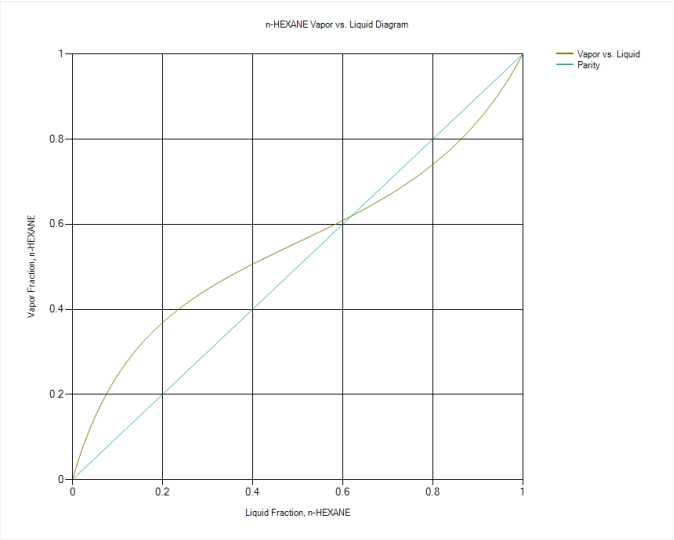


Figure 5.2. Results of stability check

Step-1: Problem definition

Input: Product type is solvent for azeotrope breaking to be used in extractive distillation
Tools: Knowledge base of solvent design
Output: Product needs, target properties and their constraints

A knowledge base, literature search and legislation details are used to determine the product needs in this case study. The main function of the solvent is to be more selective to ethyl acetate than n-hexane. In addition, it must form a totally miscible liquid with ethyl acetate and n-hexane. The solvent must not form azeotrope with neither n-hexane nor ethyl acetate. The solvent must be easy to recover and recycle and must satisfy EH&S properties. Once all needs are defined, the developed knowledge base for solvents is used to transform the product needs into target properties. The translated properties and their target property constraints are given in **Table 5.10**. The solubility parameters of the solvent should be closer to ethyl acetate (δ = 18.34 $MPa^{1/2}$) than n-hexane (δ = 14.9 $MPa^{1/2}$). The boiling point of the solvent should be higher than ethyl acetate (T_b = 350.21 K) and n-hexane (T_b = 341.88 K) in order to be easily separated from ethyl acetate after the extractive distillation is done in order to recycle the solvent back to the extractive distillation column. In order for a solvent to be attractive to ethyl acetate, the solvent should have a higher selectivity with ethyl acetate compared with n-hexane.

Table 5.9. Solvent needs, target properties and target values

Need	Property	Target value
Miscibility with the system	Solubility parameter at 298.15 K (δ), $MPa^{1/2}$	$16.34 \leq \delta \leq 18.34$
High boiling point compared to the system	Normal boiling point (T_b), K Molecular weight (M_w), g/mol	$360 \leq T_b \leq 440$ $90 \leq M_w \leq 160$
Higher selectivity to ethyl acetate than n-hexane	Selectivity	Selectivity ≥ 1.9

Step-2: Problem formulation

Input: List of target properties and their constraints from step-1.
Tools: "Azeotropic Search" options, "Advanced Search" options and ProCAMD
Output: Set of solvent candidates

The problem can be solved through:

(1) Literature search: VPPD-Lab has this option to search for azeotropic information of binary azeotropes that is collected from literatures. For the n-hexane and ethyl acetate mixture, the azeotrope is pressure maximum and homogeneous azeotrope meaning that the total pressure goes through a maximum at the constant temperature phase diagram, and therefore, the temperature goes through a minimum at the constant pressure phase diagram. In **Figure 5.3**, the experimental data at the condition that the mixture forms azeotrope is highlighted. The pressure swing distillation is recommended for this separation. The solution is shown in the "Pressure Swing Information" box. However, it also gives the solvent "n-hexyl formate" as a candidate for solvent-based extractive distillation;

(2) Database search: the advanced search option in VPPD-Lab is employed to find solvents based on search of data. The constraints are δ and Tb from **Table 5.10** are employed. The results are shown in **Figure 5.4**. A total list of 119 solvent candidates as listed in **Appendix B.1**. Note that, n-hexyl formate is also included in this list;

(3) Generate and test algorithm in ProCAMD: The third option is to use ProCAMD. The property constraints are given in terms of non-temperature pure component properties, mixture properties and azeotrope/miscibility calculations. **Figure 5.5** shows the problem specifications in ProCAMD. For the general problem specification, the search is made for acyclic hydrocarbons, ketones and esters (aromatic compounds, chlorides, dioxanes are not considered for EH&S concerns). Within each molecular class, molecular types are pre-selected to be the building blocks to generate solvent candidates. The property constraints are given in terms of non-temperature pure component properties, mixture properties and azeotrope/miscibility

calculations. The UNIFAC model is selected for the mixture specification. The feed mixture (0.49 molar fraction of n-hexane and 0.51 molar fraction of ethyl acetate) at 351.25 K and 1.0132 bar is specified. Ethyl acetate is selected as a solute that needs to be extracted with the solvent. The solvent must not also form azeotropes with ethyl acetate or n-hexane. The miscibility calculation is included to ensure that the solvent candidates are miscible with all ethyl acetate and n-hexane. **Figure 5.6** shows a screen shot from ProCAMD with the calculation results. The summary page from the screenshot indicates the following: number of solvent candidates generated, number of solvent candidates selected. If the solvent candidate is available in VPPD-Lab database, "databank" will retrieve the name of the solvent.

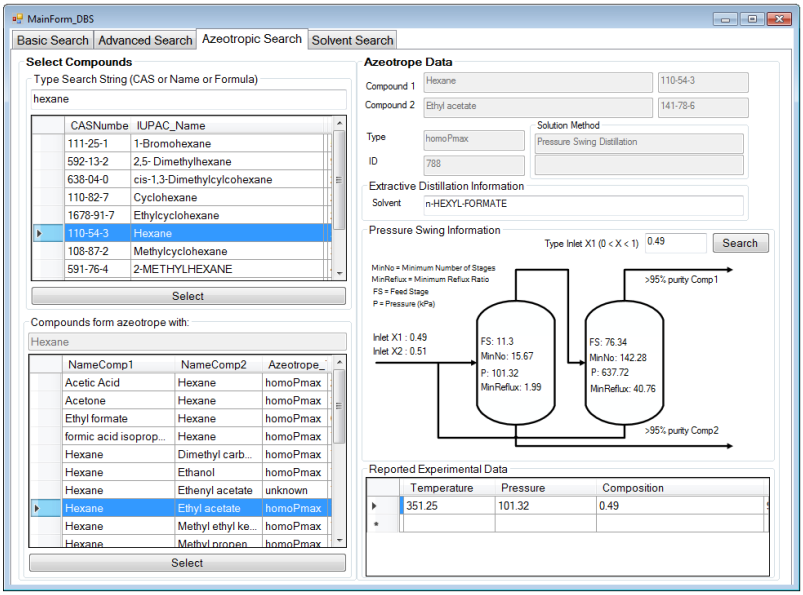


Figure 5.3. Azeotrope data of n-hexane and ethyl acetate mixture from "Azeotropic Search" template

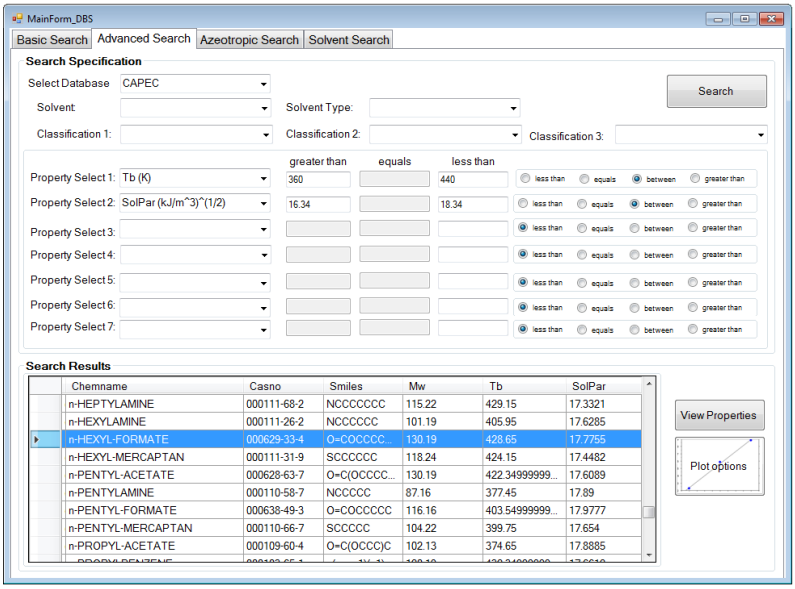


Figure 5.4. List of solvent candidates from "Advanced Search" template

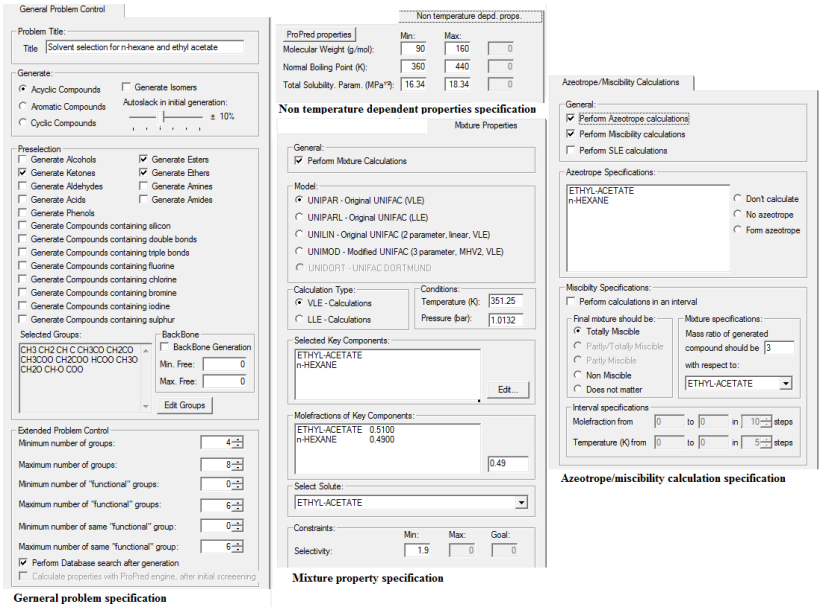


Figure 5.5. Solvent specifications in ProCAMD

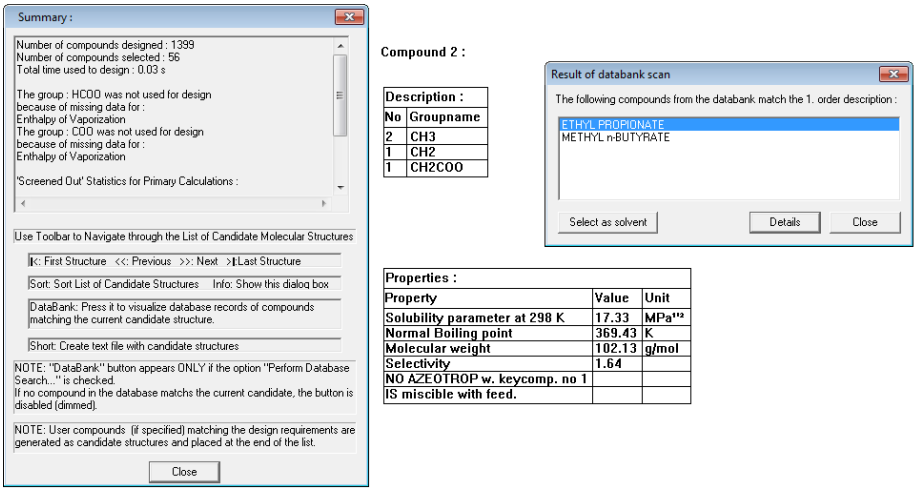


Figure 5.6. Screen shot results from ProCAMD

5.3 Blend design

Mixing or blending of two or more different chemicals is required when a single species product could not satisfy target properties. Blending fuels such as gasoline fuels, diesel fuels and diesel fuels with renewable or alternative fuels could reduce the consumption of conventional fuel, which can preserve crude oil supplies to an extent. Furthermore, adding renewable or alternative fuels could reduce the harmful exhaust emissions.

5.3.1 Design of a jet-fuel blend with alternative fuels

The objective of this case study is to minimize the conventional jet-fuel consumption by adding alternative fuels that can help to improve jet-fuel properties. Regarding the environmental sustainability, energy supply diversity and competition for energy resources for jet-fuels industries, blending alternative fuels could provide the benefits in terms of energy diversity thus reducing dependence on petroleum crude oil. In term of environmental sustainability, these biofuels could help to reduce the life cycle carbon dioxide emissions from transportation fuels because the raw materials for production of biofuels are biomass created by photosynthesis of CO_2 with water, they will have net zero combustion CO_2 emissions when burned (Hileman and Strattion, 2014).

The conventional jet-fuels comprise of a large number of hydrocarbons that are extremely difficult to be simulated for better understanding of the fuel behaviors in combustions engines. Therefore, in order to simplify the jet-fuels, a surrogate fuel that has fewer compounds and emulates certain important physical properties of the jet-fuels is chosen as the main ingredient (MI) (Agosta et al., 2004). The composition of MI is given in **Table 5.5**. The alternative fuels are chemicals produced via a hydro-treatment of bio-derived esters and fatty acids (Chuck et al., 2014) as given in **Appendix B.2**.

The jet-fuel blend design problem is developed and solved using the blend design template from the workflow shown in **Figure 4.8**.

Step-1: Problem definition

Input: Product type is the jet-fuel
Tools: Blend design template and blend design knowledge base
Output: Target properties and their constraints.

The new formulation of jet-fuel blends should have good fuel performance and meet or exceed stringent requirements for worldwide fuel handling and products standards as listed in **Table 5.10**.

Table 5.10. Product needs and their target property constraints

Need	Property	Unit	Target value
Ability to be burned	Reid vapor pressure	kPa	$RVP < 1$
Flammability	Flash point	K	$T_f > 311.15$
Engine efficiency	Higher heating value	MJ/kg	$HHV > 42.8$
	Density	kg/m3	$775 < \rho < 840$
Consistency of fuel flow	Kinematic viscosity at -20 ∘C	cSt	$V < 8$
	Melting point	K	$T_m < 226.15$
Stability	Gibbs energy of mixing	-	$d/dx\ (Gmix/RT) < 0$
Environmental impacts	CO_2 emission	kg CO_2/mile	$CO_{2E} < 24.69$
	-logLC50	mol/L	$-log(LC_{50}) < 4.726$

Step-2: Problem formulation

Input: List of target properties and their constraints from step-1.
Tools: Jet-fuel database
Output: Set of jet-fuel candidates, objective function, property constraints and equations, process/process model equations.

Potential alternative fuels are examined in this case study (Chuck et al., 2014). Pure compound properties listed in **Table 5.11** have been collected for 16 alternative fuels. Missing properties are estimated through the property toolbox. The property models for pure and mixture properties are listed in **Table 5.7** and **Table 5.8** respectively. Although blending these alternative fuels with MI is able to reduce the life cycle of CO_2 emission, it is not because the change in combustion emissions of CO_2. Therefore, in order to improve the combustion performance, CO_2 emission model is used to calculate CO_2 emission from a combustion engine when fuels are used.

Step-3: Problem solution

Input: Set of additive candidates, objective function, property constraints and equations, process/process model equations.
Tools: Solvers such as GAMS, MATLAB, etc.
Output: List of promising candidates.

The product design problem is formulated as a MINLP problem, where the fuel composition is optimized (**Eqs. (5.1)**), subject to: the linear constraints and non-linear constraints. Initially, a total of 16 chemicals are listed as feasible additives, which can formulate 120 ternary mixtures (MI + additive (1) + additive (2)). 30 alternatives are excluded due to their melting point (> 226.15 K), kinematic viscosity (>8 cSt) and $-\log LC_{50}$ (< 4.726 mol/L). Then, the remaining 90 alternative at different compositions are used to calculate miscibility in order to avoid phase separation in the engine (). None of them are removed. Subsequently, 8 alternatives are left which are evaluated through mathematical programming (to find the optimal mixture compositions) with linear and/or non-linear property constraints. The linear target property constraints are: HHV, V, CO_2 emission ρ and $-\log LC_{50}$ (see **Eqs. (5.2)**), while the non-linear constraints are: RVP (see **Eqs. (5.3)**).

$$min\ F_{obj}(x) \tag{5.1}$$

$$LB \leq \sum_{i=1}^{n} x_i P_i \leq UB \tag{5.2}$$

$$\sum_{i=1}^{n} x_i \gamma_i P_i^{sat} \leq 1 \tag{5.3}$$

where x_i is molar fraction of compound i in the mixture; UB is a upper bound value of the linear property; LB is a lower bound value of the linear property; γ_i is the activity coefficient; P_i^{sat} is the vapor pressure (kPa) of compound i at 308 K.

Finally, the most promising ternary blends with the minimum conventional jet-fuel composition are listed in **Table 5.11** with target properties values. Blending MI (50.2 %vol) with decane (15.5 %vol) and nonane (34.3 %vol) helps to reduce to consumption of MI and help to improve properties (HHV and ρ). Furthermore, it reduces CO_2 emission (6.72 % compared to MI and 9.21% compared to average jet-fuel and toxicity ($-\log LC_{50}$). Note that average jet-fuel has CO2 emission 25.37 kgCO2/mile (Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990 – 2005, 2007))

Table 5.11. Mixtures matching the target properties and their estimated property values

ID	Composition (Vol%)	HHV	V	-logLC$_{50}$	CO2E	ρ	RVP	T_f
1	MI(50.2) decane(15.5) nonane(34.3)	47	2.3	4.5	23.03	796	0.77	312.8
2	MI(56.3) decane(1.1) nonane(42.6)	47	2.25	4.47	23.04	795	0.84	311.5
3	MI(54.4) limonene(0.8) nonane(44.7)	47	2.23	4.45	23.06	795	0.86	311.2
	MI (100)	44.24	3.3	4.72	24.69	769	0.55	311.8

Step-4: Model-based verification/Experimental verification

Input: List of promising candidates from step-3.
Tools: Property toolbox and experiment toolbox.
Output: A set of promising candidates and property values calculated by rigorous property models and a set of experimental tests for product design verification.

The flash point (T_f) property model employs iteration calculation to obtain the flash point of the mixture, thus it is only used for the blends from Step-3 that have been shortlisted (see **Table 5.11**). Therefore, T_f of all blends are higher than MI and satisfy aviation Jet-A1 standards (T_f > 311.15K). Furthermore, experimental toolbox suggests experimental tests to verify V, ρ, RVP, distillation profiles and JFTOT ΔP at 260 °C to ensure that the final blends meet the aviation fuel standards based on these properties.

5.3.2 Design of a jet-fuel blend with other chemicals

In this case study, the jet-fuel main ingredient (MI) is mixed with pre-selected additives to find the best blend. The target properties are given in **Table 5.10**.

The jet-fuel blend design problem is developed and solved using the blend design template from the workflow shown in **Figure 4.8**.

Step-1: Problem definition

Input: Product type is the jet-fuel
Tools: Blend design template and blend design knowledge base
Output: Target properties and their constraints.

The new formulation of jet-fuel blends should have good fuel performance and meet or exceed stringent requirements for worldwide fuel handling and products standards as listed in **Table 5.10**.

Step-2: Problem formulation

Input: List of target properties and their constraints from step-1.
Tools: CAMD tools

Output: Set of jet-fuel candidates, objective function, property constraints and equations, process/process model equations.

A set of feasible additives is generated using ProCAMD. Thousands of chemicals are screened through the pure compound constraint of molecular weight, which is reduced to 209 chemicals based on the knowledge base and existing products as a benchmark as given in **Appendix B.3**; pure compound properties listed in **Table 5.10**. Missing properties are estimated through the property toolbox. The property models for pure and mixture properties are listed in **Table 5.6** and **Table 5.7** respectively.

Step-3: Problem solution

Input: Set of additive candidates, objective function, property constraints and equations, process/process model equations.
Tools: Solvers such as GAMS, MATLAB, etc.
Output: List of promising candidates.

The product design problem is formulated as a MINLP problem, where the fuel composition is optimized, subject to target properties. Initially, a total of 209 chemicals are listed as feasible additives, which can formulate 21,736 ternary mixtures (MI + additive (1) + additive (2)). 159 additives are excluded due to their melting point (> 226.15 K), kinematic viscosity (>8 cSt) and $-\log LC_{50}$ (< 4.726 mol/L). The remaining 50 additives generate 1,225 alternatives at different compositions. Then, 473 of the mixtures are removed because they form immiscible blends. This miscibility test is needed to avoid phase separation in the engine. Subsequently, 8 alternatives are left which are evaluated through mathematical programming (to find the optimal mixture compositions) with linear and/or non-linear property constraints. The objective function is to minimize the jet-fuel main ingredient (MI) (see **Eqs. (5.1)**). The linear target property constraints are: HHV, V, CO_2 emission, ρ and $-\log LC_{50}$ (see **Eqs. (5.2)**), while the non-linear constraint is: RVP (see **Eqs. (5.3)**). Finally, the most promising ternary blends with the minimum conventional jet-fuel composition are listed in **Table 5.12** with target properties values. Blending MI (42%vol) with decane (26%vol) and 4-methylnonane (32%vol) helps to reduce to consumption of MI and help to improve properties (HHV and ρ). Furthermore, it reduces CO_2 emission (3.2% compared to MI and 5.78% compared to average jet-fuel and toxicity ($-\log LC_{50}$).

Table 5.12. Mixtures matching the target properties and their estimated property values

ID	Composition (vol%)	RVP	HHV	P	V	CO_{2E}	$-\log LC_{50}$	T_f
1	MI(41) 2,2-dimethyloctane(30) decane(29)	0.67	47	780	2.8	24.00	4.46	312
2	MI(42) decane(26) 4-methylnonane(32)	0.57	47	783	2.7	23.90	4.65	313.2
3	MI(42) decane(26) 5-methylnonane(32)	0.57	47	782	2.7	23.94	4.65	313.1
4	MI(41) decane(52) 2,7-dimethyloctane(6)	0.52	47	777	2.8	24.07	4.68	314.8
-	MI(100)	0.55	44.24	769	3.3	24.69	4.72	311.8

Step-4: Model-based verification/Experimental verification

Input: List of promising candidates from step-3.
Tools: Property toolbox and experiment toolbox.
Output: A set of promising candidates and property values calculated by rigorous property models and a set of experimental tests for product design verification.

Flash point (T_f) property model employ an iteration calculation to obtain the flash point of the mixture, thus it is only used for the blends from Step-3 that have been shortlisted (see **Table 5.12**). Therefore, T_f of all blends are higher than MI and satisfy aviation Jet-A1 standards (T_f > 311.15K). Furthermore, experimental toolbox suggests experimental tests to verify V, ρ, RVP, distillation profiles and JFTOT ΔP at 260 °C to ensure that the final blends meet the aviation fuel standards based on these properties.

5.4 Design of a diesel blend

Tailor-made diesel blends (the blend of either diesel with different diesel additives) represent one of the most promising solutions to reduce the impact of fuel consumption on the environment while retaining or even improving the performance of diesel fuel. The aim of this case study is to design a blend containing a selected MI and additives to obtain a diesel blend that has properties better than the original MI. The composition of MI is given in **Table 5.13**.

Table 5.13 Diesel MI composition

Chemicals	Composition, vol (%)
2,2,4,4,6,8,8-Heptamethylnonane (iso-cetane)	9
Cyclo-octane	44
Toluene	1
Tetradecane	31
Dodecane	15

The diesel blend design problem is developed and solved using the blend design template from the workflow shown in **Figure 4.8**.

Step-1: Problem definition

Input: Product type is the diesel-fuel
Tools: Blend design template and blend design knowledge base
Output: Target properties and their constraints.

The new formulation of diesel blends should have good fuel performance and meet or exceed stringent requirements for worldwide fuel handling and products standards as listed in **Table 5.14**.

Table 5.14 Product needs and their target properties constraints

Need	Property	Unit	Target value
Ability to be burned	Reid vapor pressure	kPa	$RVP \leq 1.38$
Flammability	Flash point	K	$Tf \geq 325.15$
Engine efficiency	Higher heating value	MJ/kg	$HHV > 35$
	Density at 15 ∘C	kg/m3	$790 \leq \rho \leq 870$
Consistency of fuel flow	Kinematic viscosity at 40 ∘C	cSt	$1.3 \leq V \leq 4.1$
Stability	Gibbs energy of mixing	-	$d/dx\ (Gmix/RT) < 0$
Environmental impacts	-logLC50	mol/L	$-log(LC_{50}) < 5$

Step-2: Problem formulation

Input: List of target properties and their constraints from step-1.
Tools: CAMD tools
Output: Set of diesel-fuel candidates, objective function, property constraints and equations, process/process model equations.

A set of feasible additives is generated using ProCAMD. Thousands of chemicals are screened through the pure compound constraint of molecular weight, which is reduced to 29 chemicals based on the knowledge base and existing products as a benchmark as given in **Appendix B.4**; Pure compound and mixture property models are listed in **Table 5.15** and **Table 5.16** respectively. Missing properties are estimated through the property toolbox.

Table 5.15. Pure component property models

Pure component property	**Equation**
Higher heating value, HHV_i	(3.18) Yunus et al., 2014
Density, ρ_i	(3.52) and PC-SAFT (Gross and Sadowski, 2001)
Vapor pressure, P_i^{sat}	(3.53) and PC-SAFT (Gross and Sadowski, 2001)
Open cup flash point, $T_{f,i}$	(3.18) Hukkerikar et al., 2012a
Lethal concentration, $LC_{50,i}$	(3.18) Hukkerikar et al., 2012b
dynamic viscosity, η_i	(3.59) and (3.49)

Table 5.16. Mixture property models and their function

Target property	Model
Dynamic viscosity at -20 °C, η	(3.64)
Kinematic viscosity, ν	(3.64) and , $\nu = \frac{\eta}{\rho}$
Higher Heating Value, *HHV*	(3.64)
Density at 15 °C, ρ	(3.64)
Reid Vapor Pressure, *RVP*	(3.65)
Open cup flash point, T_f	(3.66)
Toxicity parameter, -log(LC_{50})	(3.64)
Distillation curve	(3.73)

Step-3: Problem solution

Input: Set of additive candidates, objective function, property constraints and equations, process/process model equations.
Tools: Solvers such as GAMS, MATLAB, etc.
Output: List of promising candidates.

The product design problem is formulated as a MINLP problem, where the fuel composition is optimized (see **Eqs. (5.1)**), subject to target properties. Initially, a total of 29 chemicals are listed as feasible additives, which can formulate 29 binary mixtures (MI + additive). 3 additives are excluded due to their heating value (> 35), kinematic viscosity ($1.3 \leq V \leq 4.1$). The remaining 26 additives generate 52 alternatives at different compositions. Then, 25 of the mixtures are removed because they form immiscible blends. Subsequently, 3 alternatives are left which are evaluated through mathematical programming to find the optimal mixture compositions which satisfy linear and non-linear property constraints. The linear target property constraints (see **Eqs. (5.2)**) are: HHV, V and -logLC_{50}, while the non-linear constraint is: RVP (see **Eqs. (5.4)**). Finally, the most promising binary blends with the minimum conventional diesel MI composition are listed in **Table 5.17** with target properties values.

$$\sum_{i=1}^{n} x_i \gamma_i P_i^{sat} \leq 1.38 \qquad (5.4)$$

Table 5.17 Mixtures matching the target properties and their estimated property values

ID	Composition (vol%)	RVP	HHV	ρ	V	-logLC_{50}	T_f
1	MI(85.81) Decane (14.19)	1.02	47	790	1.91	4.21	311.41
-	MI(100)	1.09	47	799	2.09	4.30	310.38

Step-4: Model-based verification/Experimental verification

Input: List of promising candidates from step-3.
Tools: Experiment toolbox.
Output: A set of promising candidates and property values calculated by rigorous property models and a set of experimental tests for product design verification.

The flash point (T_f) property model is applied for this verification. Experiment toolbox suggests experimental verification tests for diesel blending as listed in **Table 5.18**. The experimental tests are performed in the collaboration with TEES Gas & Fuels Research Center, Texas A&M University at Qatar. It can be noted that the results predicted by VPPD-Lab are in good agreement with the experimentally measure data. The blend satisfies all the ASTM D975 specifications and therefore is the promising candidate for further studies (engine test and emission studies).

Table 5.18 Experimental verification of diesel blends

Properties	MI (Experimental)	MI (Model)	MI + Additive (Experimental)	MI + Additive (Model)	ASTM D975 (Grade 1)
Density at 15^{0}C (g/cm^3)	0.798	0.799	0.785	0.790	-
Kinematic Viscosity at 40^{0}C (mm^2/s)	1.82	2.09	1.641	1.91	1.3 -2.4
Vapor Pressure at 37.8^{0}C (kPa)	1.10	1.09	0.9	1.02	-
Cloud Point (^{0}C)	-19.5	-	-22.1	-	-
Pour Point (^{0}C)	-18.0	-	-21.0	-	-
Flash Point (^{0}C)	40.50	37.23	41.50	38.26	Min 38
HHV (MJ/kg)	47.01	47.00	47.17	47.10	-
Cetane Index (Calculated)	-	-	48.2	-	Min 40
Distillation (K)					
IBP	-	422	427	422	-
10 %vol	-	433.5	435	429	-
50 %vol	-	480	460	458	-
90 %vol	-	514	516	515	Max 561
EBP (^{0}C)	-	521	524	522	-

5.5 Design of an insect repellent lotion

The aim is to design an insect repellent formulation for the European market (non-tropical areas). The formulation template is developed and used under the workflow in

Figure 4.9. In this product design problem, the information flow in each design step is shown in VPPD-Lab layout. The objective is to minimize the ingredient's cost (Active ingredient (AI), solvents, and additives).

Step-1: Problem definition

Input: Product type is insect repellent.
Tools: Formulation template and formulation knowledge base.
Output: User needs, target properties and constraints.

In this step, the insect repellent is selected as a product type. The user needs for the insect repellent lotion are retrieved from the knowledge base then translated in to target properties and the their constraints as shown in **Table 5.19** as well as the screen shot in **Figure 5.7**.

Table 5.19 Product needs and their target properties constraints

Need	Target Property	Unit	Target value
Effectiveness	Choice of AI	-	-
Material compatibility	Choice of solvent ,	-	-
Odour	Choice of additive	-	-
Durability	Evaporate time	second	$500 \leq$ T90 ≤ 1500
Low toxicity	Human toxicity	mol/l	$0.1 \leq -log(LC_{50}) \leq 0.6$
Stability	Hildebrand solubility parameter	$Mpa^{1/2}$	$\delta_{AI} - 3 \leq \delta \leq \delta_{AI} + 3$ $\delta_{AI} - 3 \leq \delta_{add} \leq \delta_{AI} + 3$
Good spray-ability	Kinematic viscosity molar volume	cSt l/kmol	$0 \leq V \leq 5$ $20 \leq Vm \leq 45$

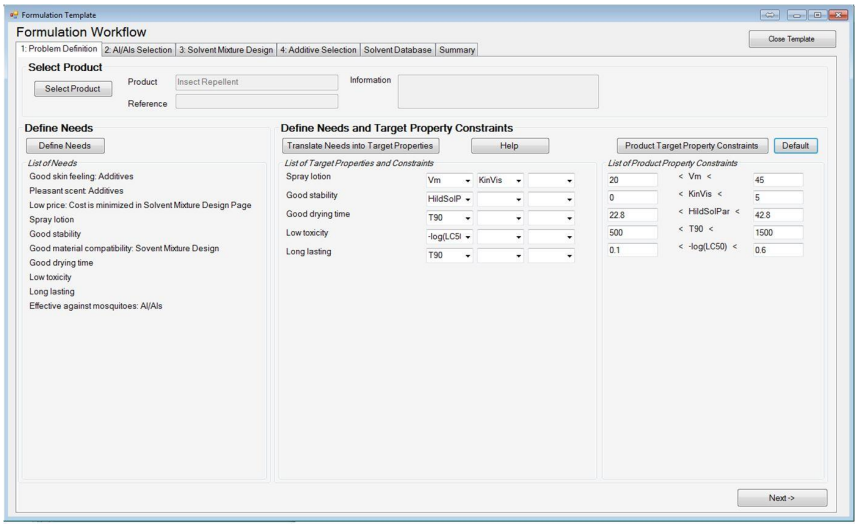


Figure 5.7. Product needs and their target property constraints of the insect repellent lotion

Step-2: AIs identification

Input: Target properties values and constraints.
Tools: Knowledge base of product functions and property toolbox.
Output: Set of AIs and properties of AIs.

An insect repellent lotion is usually constituted of the AI/AIs, with the function of repelling mosquitoes. The knowledge base of product functions is used to generate a list of active ingredient candidates with respect to the formulation product type. The list is screened based on the property constraints. In this case, the solubility parameter (HildSolPar) of active ingredient should not be less than 22.8±3 and not greater than 42.8 ±3 $MPa^{0.5}$. Picaridin is selected because it satisfies all constraints. Furthermore, it has lowest toxicity and good material compatibility compared with others in the list (see **Figure 5.8**). The properties of picaridin are retrieved from the database. The composition of AI is also suggested by the knowledge base of collected data.

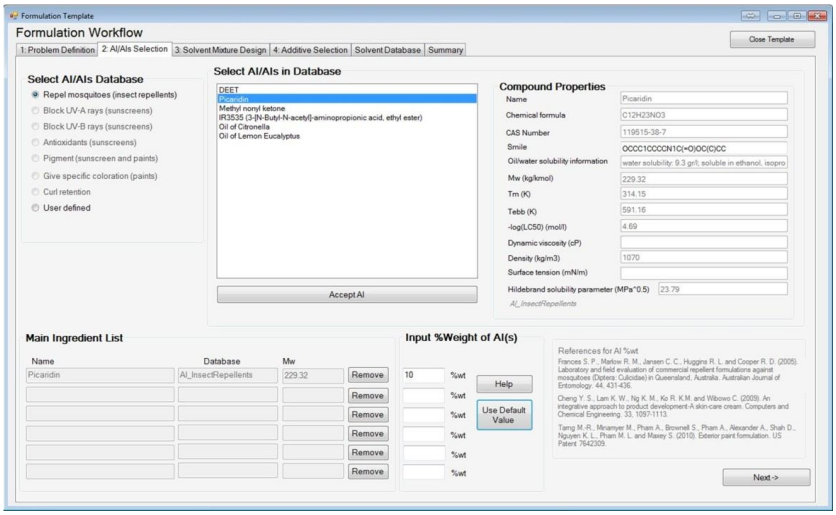


Figure 5.8. AIs identification results

Step-3: Solvent mixture design

Input: Target properties values, constraints and solvent database choices.
Tools: Property toolbox and solvers.
Output: List of promising solvent mixtures.

Since alcohols and water can be potential solvent candidates, the databases: (1) Alcohol-water soluble, (2) Alcohol-water insoluble and (3) Water are selected for solvent mixture design to be AI. The solvent mixture algorithm is then

launched. The algorithm retrieves properties of solvent candidates from the property toolbox. Lists of pure and mixture property models are given in **Table 5.20** and **Table 5.21** respectively. The product design problem is then formulates and solves the product design problem based on the product property constraints. Results are summarized in **Figure 5.9**. If the preferred performance index is the toxicity, the least toxic mixture is water + methanol. However, HildSolPa of methanol (29.3 $MPa^{0.5}$) is greater than upper bound limit (21.1 < HildSolPa < 27.1 $MPa^{0.5}$). Therefore, it has to be rejected. The most feasible mixture is water + 2-Propanol

Table 5.20. Pure property models and their function

Target property	**Model**
Cost, *c*	Database
Molar volume, *Vm*	(3.33) and PC-SAFT (Gross and Sadowski, 2001)
Vapor pressure, P_i^{sat}	(3.53) and PC-SAFT (Gross and Sadowski, 2001)
dynamic viscosity, η_i	(3.59)
Evaporate time, T90	(3.48)
Hildebrand solubility parameter,	Theoretical Equation

Table 5.21. Mixture property models and their function

Target property	**Model**
Cost, *c*	(3.64)
Kinematic viscosity, ν	(3.64) and , $\nu = \frac{\eta}{\rho}$
Molar volume, V_m	(3.64)
Hildebrand solubility parameter,	(3.64)
Toxicity parameter, -log(LC_{50})	(3.64)
Evaporate time, T90	(3.64)

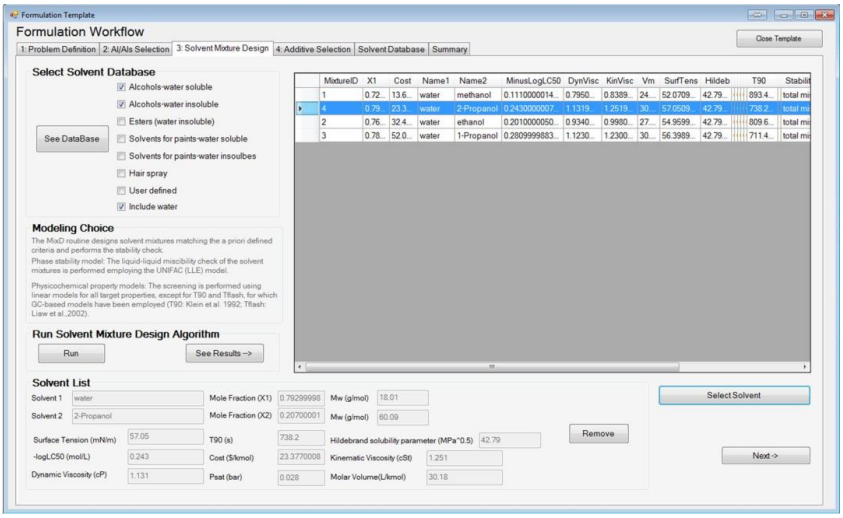


Figure 5.9. Solvent mixture design results

Step-4: Additive identification

Input: Target properties values and constraints.
Tools: Knowledge base and property toolbox.
Output: List of additives and qualities to be enhanced.

Aroma compounds can be added to give the formulation a pleasant scent. The preferred scent for the formula could be, for example, lavender. The property toolbox is used to calculate missing properties of the additive. The aroma list is screened with respect to the defined target property constraints. The additive should be alcohol soluble. Thus, the solubility of aroma compounds should be close to 2-propanol. Therefore, linalool is selected as shown in **Figure 5.10**. The composition of the selected aroma is provided by the knowledge base.

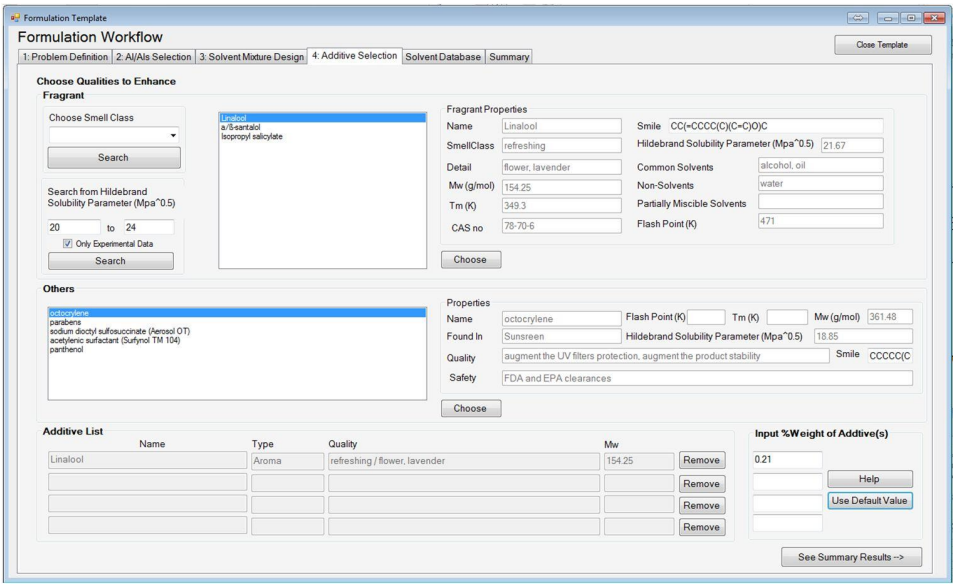


Figure 5.10. Additive identification results

Step-5: Experimental verification

Input: List of the selected formulation.
Tools: Experiment toolbox.
Output: A set of experimental tests for product design verification.

The experimental toolbox lists the experiments that need to be performed to verify the designed formulation. Note that, Conte et al (2011) verify experimentally a similar product.
The formulation template provides a summary of results together with the experimental verification list as shown in **Figure 5.11**.

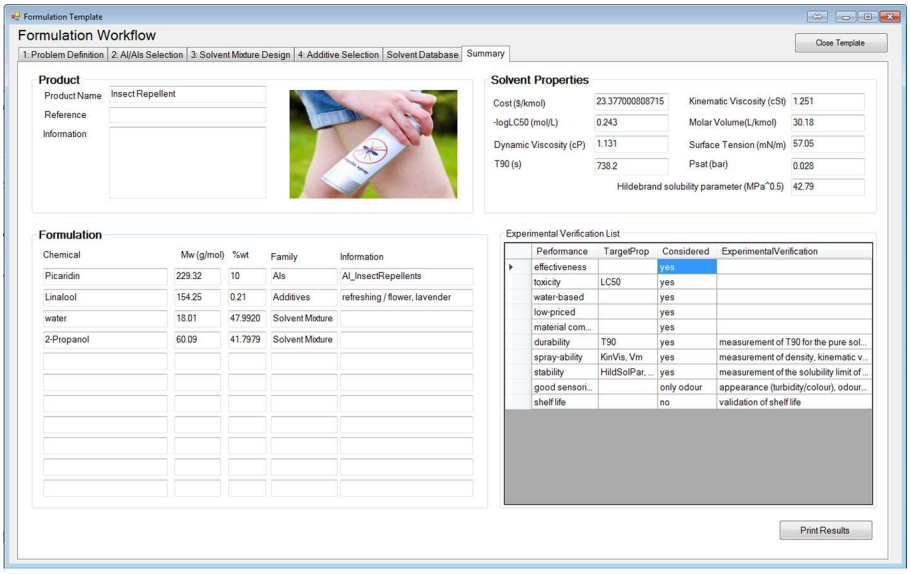


Figure 5.11. Summary results and the experimental verification list

5.6 Design of a hand-wash

Commercial hand-wash is nowadays tailor-made designed in order to satisfy a wide variety of consumer needs. The ingredients of the formulation are carefully selected in order to provide highly focused performances. This case study is adopted from Mattei et al., 2014

In this case study, the methodology for designing emulsified products implemented in the emulsion template. The main advantanges is the use of the systematic knowledge base to provide the information that are necessary for supporting the decisions and choices performed during the design.

Step-1: Problem definition

Input: Product type is hand-wash detergent
Tools: Emulsion template and emulsion knowledge base.
Output: User needs, target properties and constraints.

The consumer needs are converted into target properties using the emulsion knowledge base as shown in **Table 5.22**.

Table 5.22. Product needs and their target property constraints of an insect repellent

Need	Target Property	Unit	Target value
Foam-ability	Surface tension	mN/m	$\sigma < 40$
	Critical micelle concentration	mol/	CMC < 0.01
Non-irritability of the skin	Hansen solubility parameters	$Mpa^{1/2}$	20.4 < HanD < 24.4 7.8 < HanP < 11.8 20.49.9 < HanH < 13.9
Cleaning performances	Surface tension	mN/m	$\sigma < 55$
	Hydrophilic-lipophilic balance	-	HLB > 10
Spread-ability	Molar volume	l/mol	$30 < V_m < 150$
	Dynamic viscosity	cP	$5 < \eta < 2500$
Emulsion stability	Cloud point	K	CP > 343.15
	Krafft temperature	K	TK < 293.15
	Hydrophilic-lipophilic deviation		HLD ≠ 0
Safety	Flash point	K	Tf > 343.15
	Human toxicity	mol/m3	$-\log(LC_{50}) > 3.16$

Step-2: AIs identification

Input: Target property values and constraints.
Tools: Knowledge base of product functions and property toolbox.
Output: Set of AIs and properties of AIs.

The key consumer needs are: foam ability, non-irritability of the skin and good cleaning performances. Therefore, 2 surfactants are selected as AIs since ionic surfactants are recognized to produce high amount of foam, as opposed to non-ionic surfactants, while non-ionic surfactants are usually milder on the skin. Two databases are retrieved according to the target property constraints listed in **Figure 5.12**. A short-list of candidate ionic and non-ionic surfactants is generated. In order to select the best AIs, for the ionic surfactants, the lower surface tension and critical micelle concentration, the higher the performances; while for the non-ionic surfactants, the effectiveness is identified by Hildebrand solubility parameter. In both cases, when two or more candidate ingredients have comparable effectiveness, the cheapest solution is selected.

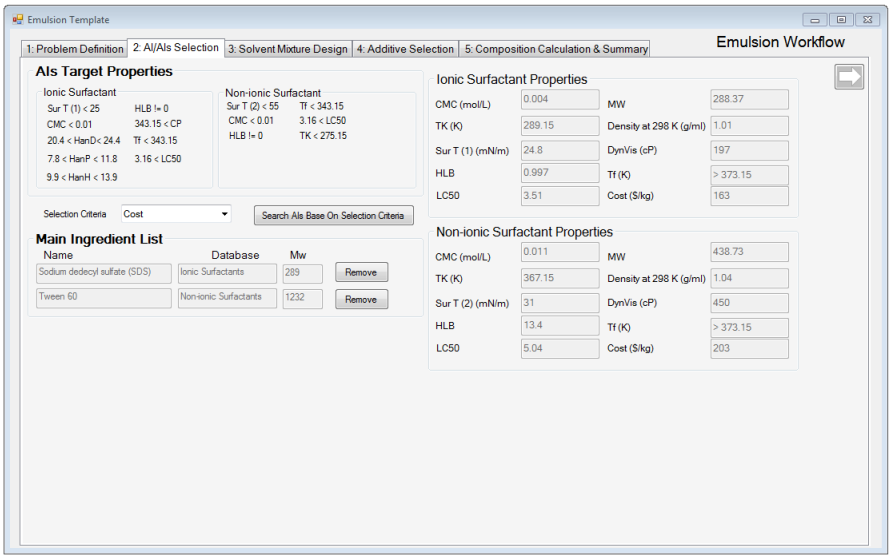


Figure 5.12. AIs identification results

Step-3: Solvent mixture design

Input: Target properties values, constraints and solvent database choices.
Tools: Property toolbox and solvers.
Output: List of promising solvent mixtures.

Hand-wash products are usually oil-in-water emulsions. In relation to the aqueous solvent phase, water is chosen as the product is directly applied on the skin, and the non-irritability of the skin is one of the needs listed in step-1. The solvent mixture algorithm is then launched. The algorithm retrieves properties of solvents candidates from the property toolbox. Note that only pure compound properties are involved in this calculation (see **Table 5.23**). The problem is formulated and solved the product design problem based on the property constraints. In order to ensure the stability of the product with respect to temperature and composition disturbances, hydrophilic-lipophilic deviation (HLD) approach is used. Negative HLD values suggest the formulation of oil-in-water emulsions, positive values suggest the formulation of water-in-oil emulsions, while HLD values in the proximity zero indicate the formation of a three-liquid-phase system, where a stable emulsion cannot be formed. The HLD values of each AI candidates are calculated. In this step, the HLD value of the product is calculated, to check if the addition of active ingredients and additives has influenced the stability of the product as an emulsion. A list of the most promising aqueous solvent phase and organic solvent phase are selected as shown in **Figure 5.13**.

Table 5.23. Pure component property models

Target Property	Equation
Hansen solubility parameters (HanD, HanP, HanH)	(3.18)
Molar volume	(3.33) and PC-SAFT (Gross and Sadowski, 2001)
Dynamic viscosity	(3.49)
Flash point	(3.18)
Human toxicity	(3.18)
Cost, C	Database

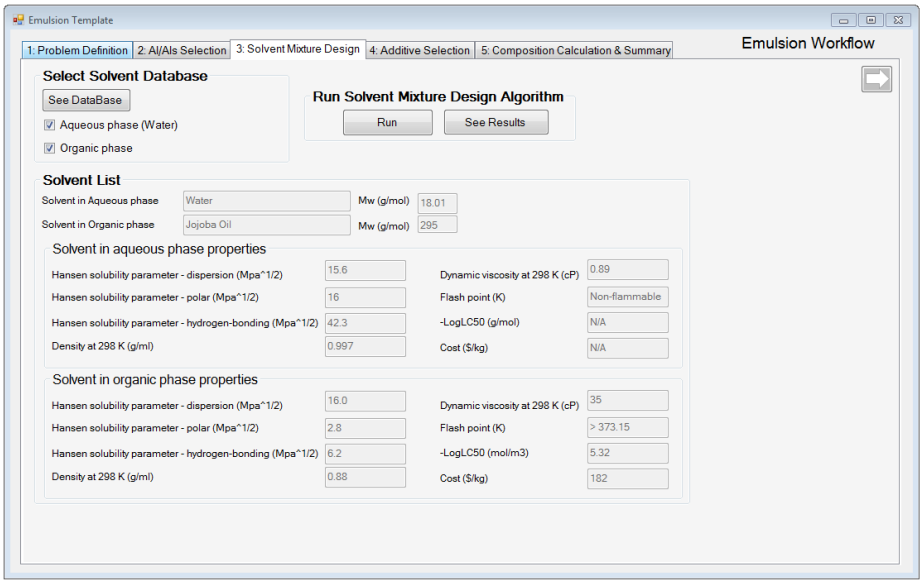


Figure 5.13. Solvent design results

Step-4: Additive identification

Input: Target properties values and constraints.
Tools: Knowledge base, additive database and property toolbox.
Output: List of additives and qualities to be enhanced.

Additives are needed to fulfill the secondary consumer needs: an aroma, a colorant, a co-surfactant, an emollient and a preservative. Note that anti-bacterial agents are not considered in this case study as the Food and Drug Administration (FDA) has recently expressed concerns over anti-microbial agents, as they are currently under investigation for potential carcinogen effects (Mattei et al., 2014). Five databases are used based on the emulsion knowledge base. The criteria for selection of the promising additives is the cost. The cheapest ingredients satisfying constraints set in step-1 are selected as shown in **Figure 5.14**.

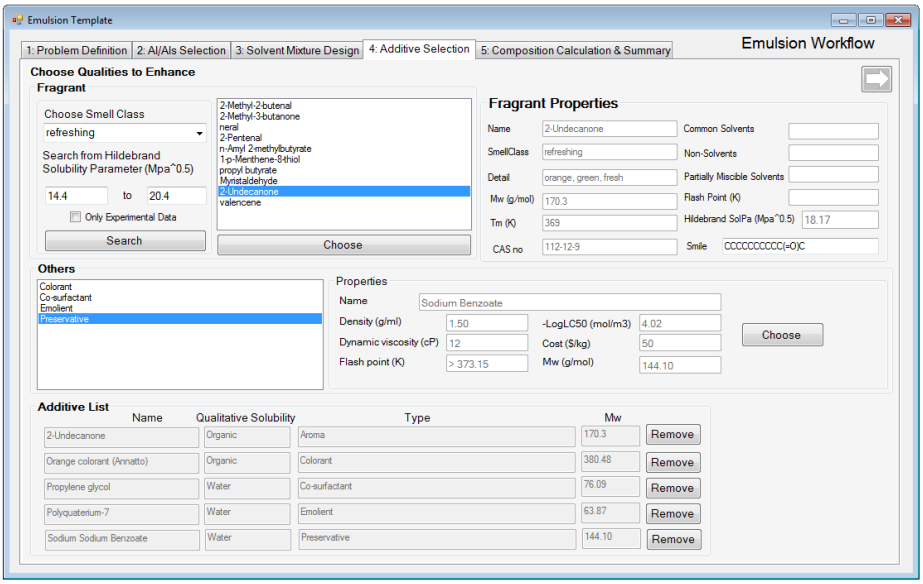


Figure 5.14. Additive identification results

Step-5: Composition calculation

Input: List of AIs, solvents and additives.
Tools: Knowledge base and property toolbox.
Output: List of the compounds and their composition.

Once all appropriate ingredients have been chosen, the composition calculator is used to determine the overall composition of the product. In this step, the solubility of the different ingredients in the two solvents is quantified with UNIFAC-based calculations. The *a priori* defined target property constraints are considered. The knowledge base is used to set feasible composition ranges of ingredients since some of them are known to be effective only in a certain range of compositions. Finally, the emulsified product is determined by minimizing the total cost. **Figure 5.15** summarizes the formulation in the emulsion form together with an experimental verification list for the hand-wash detergent.

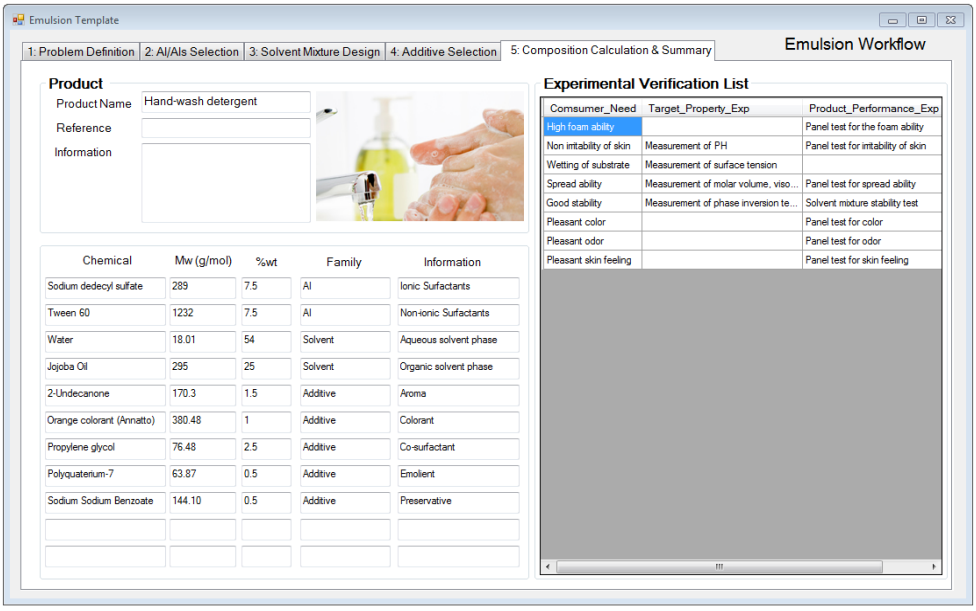


Figure 5.15. Summary results and the experimental verification list

Step-6: Experimental verification

The aim of this step is to validate and/or refine formulations by means of experiments. Note that in this case study, no experimental work has been performed. However, through the use of knowledge base, a list of experimental verification tests that could be performed for of the hand-wash detergent is shown in **Figure 5.15**. Mattei et al. (2014) performed experimentally verification tests on a similar product.

5.7 Design of microcapsules for controlled release of pesticides

In the agrochemical industry, a use of a controlled release device to deliver pesticides to crops has numerous advantages: from optimized delivery of pesticides (AIs), to reduction of hazards to humans and environment. That is, the amount of pesticide used on the field can be reduced and also the safety level during its use can be improved. The device basically consists of a pesticide as an active ingredient (AI) that is encapsulated within a polymer membrane which controls amount of the AI to be diffused out in to water (the most common release medium in the agrochemical field).

The device template is developed and used under the workflow in **Figure 4.11**.

This case study highlights the advantages of the device template in terms of software architecture (handle the needs of different model interfaces), data-flow, reliable of the

use of the template in the design of new products, finally, the design steps to achieve the end-use property targets. However, another case might be the design of microcapsules where given one active ingredient, the donor and the release medium, it is desired to develop a product having specific release behavior consisting of: 10% of active ingredient release in 3 hours (as lower limit) and 50% of active ingredient release in 1 hour (as upper limit). The active ingredient to be released is known in advance, therefore, concerns the finding of the polymeric material able to achieve the target behavior.

Step-1: Problem definition

Input: Product type and list of key ingredients (AI, candidates for donor medium, release medium, polymer membrane).
Tools: Knowledge base.
Output: Product performance constraints and list of physicochemical phenomena models.

The product to be designed is the device is a microcapsule controlled device for release of a pesticide. The product structure is shown in **Figure 4.9**. The permethrin is selected as an AI to be dissolve in a donor medium and released from poly butyl-methacrylate that forms the microcapsule to water (release medium). It is designed to release 90 % of the permethrin to water in 3 hours. The design problem, therefore, concerns the finding of the donor medium that is able to achieve the target behavior. Multiscale models are considered. Through the use of knowledge base, 3 different scales are identified: Nano-scale (calculation of diffusion coefficient), micro-scale (release of AI in presented microcapsules) and meso-scale (normal distribution of microcapsules). These models, are implemented in VPPD-Lab property obtained from Shirley et al. (2005) and Morales-Rodriguez and Gani (2009).

In order to simplify the models, the following assumptions are made:

- Diffusion occurs through a film that is thin enough so that the diffusion can be considered one dimensional.
- Initial concentration of AI is equal for all the microcapsules.
- Diffusion coefficient is independent of concentration.
- Concentration can be affected for the diffusivity of the active ingredient into the polymer, and also, due to the partition coefficient between the wall of the microcapsule (polymeric membrane) and the donor and the receiver (release medium).
- Isothermal condition during the controlled release.
- Non-constant activity source.
- The model is applicable to systems where the AI is available in solution below the solubility limit.

Step-2: Calculation of primary and secondary properties

Input: Product performance constraints, list of physicochemical phenomena models.
Tools: Knowledge base and property toolbox.
Output: Set of product performances variables.

VPPD-Lab performs the meso-scale calculation and values of variables needed in the lower scale are transferred to the micro-scale where the estimation of the release is done. Afterwards, information for each set of microcapsules with the different sizes are calculated and returned to the meso-scale where the total estimation of the microcapsule based controlled release is performed. The Data-flow for the microcapsule controlled release is shown in **Figure 3.19**. The first step involving the calculation of the diffusivity coefficient between AI and the polymer membrane is calculated using the models from Muro-Suñé (2005). The calculation of some missing properties such as critical volume, molecular weight and glass temperature are calculated via the property toolbox to fill the gap. The next step involves the calculation of partition coefficients for the donor medium polymeric membrane wall and the release medium polymeric membrane wall.

Step-3: Calculation of product performances

Input: Set of model variables.
Tools: Knowledge base and property toolbox.
Output: Product performance results.

Once all the necessary data and information have been retrieved, the product performance (controlled release of AI) is selected, and all collected information and/or calculated information (such as, diffusion coefficient, partition coefficient, etc.) are retrieved as shown in **Figure 5.16**.

Simulation results from the solution of the controlled release model using different donor mediums are shown in **Figure 5.17**. Given the input data and geometry chosen for the microcapsule, it is possible to archive 95% release from the capsule in 3 hours by using n-hexane as the donor medium. Therefore, the device template is able to help to screen criteria of product candidates by calculation of the product properties and performances and test if the candidate satisfies the target behavior.

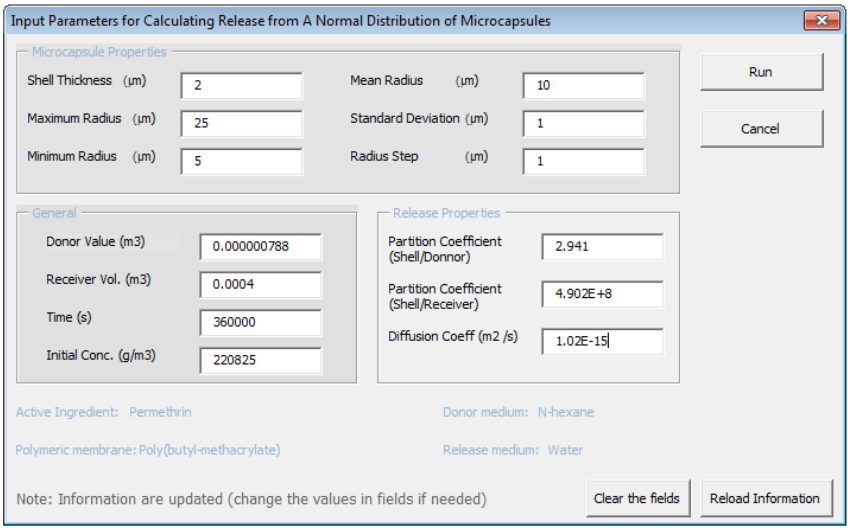


Figure 5.16. Model data for the microcapsule for controlled release

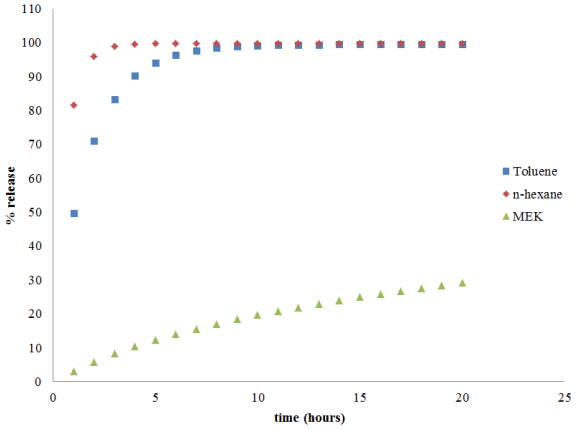


Figure 5.17. Controlled release behavior of permethrin

Step-4: Experimental verification

In order to improve the calculation results, the experimental toolbox suggests to measure the radius and wall thickness of the microcapsules. The measurement of % release can be done to verify the accuracy of the models. Furthermore, after the controlled release behavior has been archived, it is also necessary to calculate how much the AI (permethrin) is taken up by the cuticle and the plant , this is be done through the product analysis template. Through the use of the product analysis template, the uptake of permethrin can be simulated.

5.8 Other application examples

Many more application examples of VPPD-Lab can be found in **Table 5.24.** The property data and models related to the design of single molecule products such as solvents, refrigerants and active ingredients are available in VPPD-Lab. ProCAMD and SolventPro can be used to generate product candidates with respect to target properties such as boiling point, melting point, separation factor, solvent loss for designing of solvents. The molecular design template can be used to formulate mathematical programing problem to design a refrigerant by considering product and process constraints (Cignitti et al., 2015). Property data and models for pure compounds and mixtures in VPPD-Lab can be used to design blended products especially for fuels blending where the common properties are higher heating value, liquid viscosity, liquid density, liquid heat capacity, vapor pressure, liquid phase stability and distillation temperature. For formulation design and emulsion design, the databases for active ingredients such as insect repellent agents, surfactants, sun protection agents, aromas, paints that are in VPPD-Lab together with design algorithms can be used to design insect repellent lotions, sunscreen lotion, paints, detergents and various skin-care cream. The models to predict product performance are the core of the device design. In VPPD-Lab, the uptake of pesticides from water droplets to leaves and microcapsule controlled release of active ingredients are implemented. For product analysis, the suit of property models in VPPD-Lab can be used to calculate properties of pure chemical compounds and mixtures. The possibility of separation of various mixtures can be analyzed to find the appropriate methods. For example, the vapor-liquid phase equilibria of acetone and chloroform can be done to identify the azeotrope. The solvent for extraction of acetone from chloroform can be verified by calculation of ternary diagram to ensure that no azeotrope with either acetone or chloroform is formed and to identify the separation boundaries.

Table 5.24. Product design and analysis problems solved through VPPD-Lab

Product design problem	Product	Reference
Molecular design		
- Solvent mixture design	Solvent substitutes for separation of acetic acid from water, Solvent-Anti solvent design for Ibuprofen,	Hostrup et al.,1999; Karunanithi et al, 2004; Karunanithi et al, 2005; Mitrofanov et al., 2012;
- Chemical replacement	Prediction of multicomponent	Harper et al.,2000; Gani et al.,
- Polymer design	diffusion , Solvent replacement for	2005; Satyanarayana et al.,
- Refrigerant design	multistep organic synthesis, polymers and refrigerants	2009; Cignitti et al., 2015; Churi and Achenie, 1996;
Blend design		
		Yunus et al., 2012; Yunus et al.,
	Gasolines, lubricants, diesels and	2013; Phoon et al., 2015;
Formulation design	jet-fuels	Kalakul et al., 2015
	Insect repellent lotion, sunscreen	Cheng et al., 2009; Conte et al.,
Emulsion design	lotion, paint formulation and skin-care cream	2011; Conte et al., 2012;
		Mattei et al., 2013; Mattei et al.,
Device design	Tank cleaning detergent and hand-wash detergent	2014;
		Morales-Rodriguez et al., 2009;
	Uptake of pesticides from water droplets to leaves, microcapsule	Morales-Rodriguez et al., 2011; Teixeira et al., 2012;
Product analysis	controlled release of active	
- Product separation	ingredients	
- Miscibility calculation		
- Phase equilibria		Hostrup et al.,1999; Cunico et
calculation	Acetone/ chloroform separation, lipids separation, ionic liquids separation, solvents separation, solvent stability tests, solvent evaporation test, uptake of active ingredient, microcapsule controlled release of active ingredients, VLE/SLE/LLE calculations	al., 2013; Cunico et al., 2014; Morales-Rodriguez et al., 2009;

6 CONCLUSION AND FUTURE WORK

A computer-aided framework for design of chemical products has been developed and used as the architecture for developing the VPPD-Lab software.

Achievements

In order to design CPs, the representation of each CP is needed in order to understand the product and appropriate property models that can be reliably used. The collection of product representation has been established (see section 3.1 in Chapter 3). Since the reliability of CPDs depends on the how to identify the needs for a specific product, and relating these needs to physicochemical properties. The knowledge base that store product needs and experimental verification of each CP has been created (see section 3.2.1 in Chapter 3). Property models to: generate feasible product candidates and verify if the candidate satisfies desired target properties and/or product performance. The algorithms for generation of product candidates of each CP have been collected (see section 3.1 in Chapter 3). The property models that are available in literatures as well as the new liquid viscosity model of complex esters and all the needed interaction parameters are collected and stored in the property model library (see section 3.3 in Chapter 3). The CPD methodologies that provide workflow and dataflow to employ computer-aided tools to screen product candidates in order to identify the promising candidates are collected (see section 3.4 in Chapter 3). Since the solution of CPD problems required different methodologies and tools such as database, property models, design algorithms, solvers, product design tools (such as ProCAMD and SolventPro). The methodologies and tools are integrated into a systematic framework for chemical product design and evaluation that provides the work-flows and dataflow of the methodologies through the product design ontology developed to represent the associated knowledge (see section 3.6 in Chapter 3). The new computer-aided tool for chemical product design and evaluation has been created (Kalakul et al., 2015) in order to utilized the systematic framework to solve a wide range of CPD problems in an easy and efficient manner by providing the necessary methods, tools, and references (see Chapter 4). Product design templates have been generated to provide corresponding product design/evaluation workflow, the associated data-flow, tools, models, and calculation algorithms for each type of CPs. This way, a product designer is able to fastly change available options (such as chemical compounds, property models, product target property constraints) allowed by each product design template based on the product design goals (see section 4.2 in Chapter 4). The software architecture and the product design templates are able to handle the complexity of product design case studies and analysis problem case studies, in terms of use of models, calculation

algorithms, use of databases and the various problem specific solution strategies (see Chapter 5). Therefore, the main issues and needs that were to be achieved in this work are listed as:

- Collection of methods and tools for solution of CPD problems;
- Collection and implementation of property models (also development if necessary);
- Creation of systematic framework for CPD utilizing the product design templates and associated tools;
- Creation of product design templates, based on product types and design methodologies, which includes the models, algorithms, tools, databases and knowledge base;
- Development of case studies to test and validate the framework and the VPPD-Lab software.

All above issues and needs have been achieved in this work.

The use of the templates and the VPPD-Lab software is highlighted through case studies. The application of the product analysis template is highlighted through the property estimation of pure and mixture properties of the gasoline surrogate and a jet-fuel surrogate. The property estimation results from the estimation are closed to the experimental test results. Therefore, property models employed in VPPD-Lab are predictive. In addition, the application of the product design template is highlighted through case studies involving mixture/blend design of a jet-fuel and a lubricant as blended liquid products and insect repellent lotion as a formulation product. The product design template is able to handle the large mixed-integer non-linear problem formulated to design the three products. It helps to reduce the search space and provides promising chemical candidates that are competitive, and environmentally feasible, making it more flexible and capable of solving a wide range of product design problems. Therefore, VPPD-Lab enhances the future development of chemical product design as huge amounts of data, models, knowledge, methodologies and algorithms are integrated and managed in a systematic and efficient way, increasing the possibility to capture past experiences and provide better guidelines for future chemical products.

Challenges and future work

Despite the advances made in this PhD thesis, with the currently available methods and tools only a small percentage of chemical product design problems can be solved. Much work and concerted efforts are needed in the area of property modeling and their integration with data and design tools that incorporate data-models in multidisciplinary solution approaches to cover a wider range of chemicals based products of significance. Therefore, future work should address:

- Modeling issues: more reliable property models for prediction of product properties and functions are needed for many large molecular products such as pharmaceutical ingredients, proteins, biomass, membranes and cosmetics.

- Devices: while the design of devices is still based on generate and test approach, the development of model-based optimization methods to find the optimal sequence of process steps.
- Product-process design: integration of product-process design modeling is able to reduce the search space for finding promising chemical products and processes. In order to archive, process models together with the needed property models should be developed;
- Life cycle assessment (LCA) and sustainability analysis: integration of LCA aspects as well as product-process sustainability is challenging since global warming brought about by widespread environmental pollutions, resource depletions, rising human population, and multiple threats to food, water and energy securities require a paradigm shift in engineering thinking and ways to find and test solutions. Chemical product-process for the future should take these issues into account.

APPENDIX A

SPEED Lipids Database

Table A.1 List of compound in the database

No.	Name	Class	Formula
1	Alpha-Carotene	Carotenoids	C40H56
2	Beta-Carotene	Carotenoids	C40H56
3	Delta-Carotene	Carotenoids	C40H56
4	Epsilon-Carotene	Carotenoids	C40H56
5	Gamma-Carotene	Carotenoids	C40H56
6	Lutein	Carotenoids	C40H56O2
7	Lycopene	Carotenoids	C40H56
8	Zeaxanthin	Carotenoids	C40H56O2
9	1,2-didecanoyl-sn-glycerol	Diglycerides	C23H44O5
10	1-decanoyl-2dodecanoyl-sn-glycerol	Diglycerides	C25H48O5
11	1-hexanoyl-2octanoyl-sn-glycerol	Diglycerides	C17H32O5
12	1-decanoyl-2octadecenoyl-sn-glycerol	Diglycerides	C31H58O5
13	1,2-dioctanoyl-sn-glycerol	Diglycerides	C19H36O5
14	1-octanoyl-2dodecanoyl-sn-glycerol	Diglycerides	C23H44O5
15	1-octanoyl-2octadecenoyl-sn-glycerol	Diglycerides	C29H54O5
16	1-octadecadienoyl-2eicosanoyl-sn-glycerol	Diglycerides	C41H76O5
17	1-octadecadienoyl-2docosanoyl-sn-glycerol	Diglycerides	C43H80O5
18	1-octadecadienoyl-2docosenoyl-sn-glycerol	Diglycerides	C43H78O5
19	1-octadecadienoyl-2eicosenoyl-sn-glycerol	Diglycerides	C41H74O5
20	1,2-dioctadecadienoyl-sn-glycerol	Diglycerides	C39H68O5
21	1-octadecadienoyl-2octadecatrienoyl-sn-glycerol	Diglycerides	C39H66O5
22	1,2-didodecanoyl-sn-glycerol	Diglycerides	C27H52O5
23	1-dodecanoyl-2tetradecanoyl-sn-glycerol	Diglycerides	C29H56O5
24	1,2-dioctadecatrienoyl-sn-glycerol	Diglycerides	C39H64O5
25	1-dodecanoyl-2octadecenoyl-sn-glycerol	Diglycerides	C33H62O5
26	1-dodecanoyl-2hexadecanoyl-sn-glycerol	Diglycerides	C31H60O5
27	1-dodecanoyl-2octadecanoyl-sn-glycerol	Diglycerides	C33H64O5
28	1-heptadecanoyl-2octadecenoyl-sn-glycerol	Diglycerides	C38H72O5
29	1-tetradecanoyl-2octadecadienoyl-sn-glycerol	Diglycerides	C35H64O5
30	1,2-ditetradecanoyl-sn-glycerol	Diglycerides	C31H60O5
31	1-tetradecanoyl-2octadecenoyl-sn-glycerol	Diglycerides	C35H66O5
32	1-heptadecenoyl-2octadecenoyl-sn-glycerol	Diglycerides	C38H70O5

33	1-tetradecanoyl-2hexadecanoyl-sn-glycerol	Diglycerides	C33H64O5
34	1-octadecenoyl-2eicosanoyl-sn-glycerol	Diglycerides	C41H78O5
35	1-octadecenoyl-2docosanoyl-sn-glycerol	Diglycerides	C43H82O5
36	1-octadecenoyl-2docosenoyl-sn-glycerol	Diglycerides	C43H80O5
37	1-octadecenoyl-2eicosenoyl-sn-glycerol	Diglycerides	C41H76O5
38	1-octadecenoyl-2octadecadienoyl-sn-glycerol	Diglycerides	C39H70O5
39	1-octadecenoyl-2octadecatrienoyl-sn-glycerol	Diglycerides	C39H68O5
40	1,2-dioctadecenoyl-sn-glycerol	Diglycerides	C39H72O5
41	1-hexadecanoyl-2eicosanoyl-sn-glycerol	Diglycerides	C39H76O5
42	1-hexadecanoyl-2docosanoyl-sn-glycerol	Diglycerides	C41H80O5
43	1-hexadecanoyl-2octadecadienoyl-sn-glycerol	Diglycerides	C37H68O5
44	1-hexadecanoyl-2octadecatrienoyl-sn-glycerol	Diglycerides	C37H66O5
45	1-hexadecanoyl-2octadecenoyl-sn-glycerol	Diglycerides	C37H70O5
46	1,2-dihexadecanoyl-sn-glycerol	Diglycerides	C35H68O5
47	1-hexadecanoyl-2octadecanoyl-sn-glycerol	Diglycerides	C37H72O5
48	1-octadecanoyl-2octadecenoyl-sn-glycerol	Diglycerides	C39H74O5
49	1,2-dioctadecanoyl-sn-glycerol	Diglycerides	C39H76O5
50	decanoic acid , ethyl ester	Ethyl esters	C12H24O2
51	undecanoic acid, ethyl ester	Ethyl esters	C13H26O2
52	dodecanoic acid, ethyl ester	Ethyl esters	C14H28O2
53	tridecanoic acid, ethyl ester	Ethyl esters	C15H30O2
54	tetradecenoic acid, ethyl ester	Ethyl esters	C16H30O2
55	tetradecanoic acid, ethyl ester	Ethyl esters	C16H32O2
56	pentadecanoic acid, ethyl ester	Ethyl esters	C17H34O2
57	hexadecadienoic acid, ethyl ester	Ethyl esters	C18H32O2
58	hexadecenoic acid, ethyl ester	Ethyl esters	C18H34O2
59	hexadecanoic acid, ethyl ester	Ethyl esters	C18H36O2
60	heptadecenoic acid, ethyl ester	Ethyl esters	C19H36O2
61	heptadecanoic acid, ethyl ester	Ethyl esters	C19H38O2
62	octadecatrienoic acid, ethyl ester	Ethyl esters	C20H34O2
63	octadecadienoic acid, ethyl ester	Ethyl esters	C20H36O2
64	octadecenoic acid, ethyl ester	Ethyl esters	C20H38O2
65	octadecanoic acid, ethyl ester	Ethyl esters	C20H40O2
66	eicosatetraenoic acid, ethyl ester	Ethyl esters	C22H36O2
67	eicosadienoic acid, ethyl ester	Ethyl esters	C22H40O2
68	eicosenoic acid, ethyl ester	Ethyl esters	C22H42O2
69	eicosanoic acid, ethyl ester	Ethyl esters	C22H44O2
70	docosenoic acid, ethyl ester	Ethyl esters	C24H46O2
71	docosanoic acid, ethyl ester	Ethyl esters	C24H48O2
72	tetracosenoic acid, ethyl ester	Ethyl esters	C26H50O2
73	tetracosanoic acid, ethyl ester	Ethyl esters	C26H52O2
74	hexanoic acid, ethyl ester	Ethyl esters	C8H16O2
75	heptanoic acid, ethyl ester	Ethyl esters	C9H18O2
76	nonanoic acid, ethyl ester	Ethyl esters	C11H22O2
77	octanoic acid, ethyl ester	Ethyl esters	C10H20O2
78	hexanoic acid,2-ethylhexyl ester	Ethylhexyl esters	C14H28O2
79	octanoic acid,2-ethylhexyl ester	Ethylhexyl esters	C16H32O2
80	decanoic acid ,2-ethylhexyl ester	Ethylhexyl esters	C18H36O2

81	dodecanoic acid,2-ethylhexyl ester	Ethylhexyl esters	C20H40O2
82	tetradecanoic acid,2-ethylhexyl ester	Ethylhexyl esters	C22H44O2
83	hexadecanoic acid,2-ethylhexyl ester	Ethylhexyl esters	C24H48O2
84	octadecadienoic acid, 2-ethylhexyl ester	Ethylhexyl esters	C26H48O2
85	octadecenoic acid, 2-ethylhexyl ester	Ethylhexyl esters	C26H50O2
86	octadecanoic acid, 2-ethylhexyl ester	Ethylhexyl esters	C26H52O2
87	docosanoic acid	Fatty acids	C22H44O2
88	eicosanoic acid	Fatty acids	C20H40O2
89	decanoic acid	Fatty acids	C10H20O2
90	eicosapentaenoic acid	Fatty acids	C20H30O2
91	eicosatetraenoic acid	Fatty acids	C20H32O2
92	docosenoic acid	Fatty acids	C22H42O2
93	eicosenoic acid	Fatty acids	C20H38O2
94	eicosadienoic acid	Fatty acids	C20H36O2
95	hexanoic acid	Fatty acids	C6H12O2
96	heptanoic acid	Fatty acids	C7H14O2
97	dodecanoic acid	Fatty acids	C12H24O2
98	tetracosanoic acid	Fatty acids	C24H48O2
99	octadecadienoic acid	Fatty acids	C18H32O2
100	octadecatrienoic acid	Fatty acids	C18H30O2
101	heptadecanoic acid	Fatty acids	C17H34O2
102	heptadecenoic acid	Fatty acids	C17H32O2
103	tetradecenoic acid	Fatty acids	C14H26O2
104	tetracosenoic acid	Fatty acids	C24H46O2
105	octanoic acid	Fatty acids	C8H16O2
106	octadecenoic acid	Fatty acids	C18H34O2
107	hexadecanoic acid	Fatty acids	C16H32O2
108	hexadecenoic acid	Fatty acids	C16H30O2
109	pentadecanoic acid	Fatty acids	C15H30O2
110	(9Z,12R)-12-Hydroxyoctadec-9-enoic acid	Fatty acids	C18H34O3
111	octadecanoic acid	Fatty acids	C18H36O2
112	tridecanoic acid	Fatty acids	C13H26O2
113	undecanoic acid	Fatty acids	C11H22O2
114	tetradecanoic acid	Fatty acids	C14H28O2
115	nonanoic acid	Fatty acids	C9H18O2
116	1-hexanol	Fatty alcohol	C6H14O
117	1-octanol	Fatty alcohol	C8H18O
118	1-decanol	Fatty alcohol	C10H22O
119	1-heptanol	Fatty alcohol	C7H16O
120	1-dodecanol	Fatty alcohol	C12H26O
121	1-tetradecanol	Fatty alcohol	C14H30O
122	1-hexadecanol	Fatty alcohol	C16H34O
123	1-octadecanol	Fatty alcohol	C18H38O
124	1-eicosanol	Fatty alcohol	C20H42O
125	1-docosanol	Fatty alcohol	C22H46O
126	octadecenoic acid, isopropyl ester	Isopropyl esters	C21H40O2
127	hexadecanoic acid, isopropyl ester	Isopropyl esters	C19H38O2
128	octadecanoic acid, isopropyl ester	Isopropyl esters	C21H42O2

129	nonanoic acid, methyl ester	Methyl esters	C10H20O2
130	decanoic acid , methyl ester	Methyl esters	C11H22O2
131	undecanoic acid, methyl ester	Methyl esters	C12H24O2
132	dodecanoic acid, methyl ester	Methyl esters	C13H26O2
133	tridecanoic acid, methyl ester	Methyl esters	C14H28O2
134	tetradecenoic acid, methyl ester	Methyl esters	C15H28O2
135	tetradecanoic acid, methyl ester	Methyl esters	C15H30O2
136	pentadecanoic acid, methyl ester	Methyl esters	C16H32O2
137	hexadecadienoic acid, methyl ester	Methyl esters	C17H30O2
138	hexadecenoic acid, methyl ester	Methyl esters	C17H32O2
139	hexadecanoic acid, methyl ester	Methyl esters	C17H34O2
140	heptadecenoic acid, methyl ester	Methyl esters	C18H34O2
141	heptadecanoic acid, methyl ester	Methyl esters	C18H36O2
142	octadecatrienoic acid, methyl ester	Methyl esters	C19H32O2
143	octadecadienoic acid, methyl ester	Methyl esters	C19H34O2
144	octadecenoic acid, methyl ester	Methyl esters	C19H36O2
145	octadecanoic acid, methyl ester	Methyl esters	C19H38O2
146	eicosatetraenoic acid, methyl ester	Methyl esters	C21H34O2
147	eicosadienoic acid, methyl ester	Methyl esters	C21H38O2
148	eicosenoic acid, methyl ester	Methyl esters	C21H40O2
149	eicosanoic acid, methyl ester	Methyl esters	C21H42O2
150	docosenoic acid, methyl ester	Methyl esters	C23H44O2
151	docosanoic acid, methyl ester	Methyl esters	C23H46O2
152	tetracosenoic acid, methyl ester	Methyl esters	C25H48O2
153	tetracosanoic acid, methyl ester	Methyl esters	C25H50O2
154	hexanoic acid, methyl ester	Methyl esters	C7H14O2
155	heptanoic acid, methyl ester	Methyl esters	C8H16O2
156	octanoic acid, methyl ester	Methyl esters	C9H18O2
157	1-eicosanoyl-sn-glycerol	Monoglycerides	C23H46O4
158	1-docosanoyl-sn-glycerol	Monoglycerides	C25H50O4
159	1-decanoyl-sn-glycerol	Monoglycerides	C13H26O4
160	1-octanoyl-sn-glycerol	Monoglycerides	C11H22O4
161	1-docosenoyl-sn-glycerol	Monoglycerides	C25H48O4
162	1-eicosenoyl-sn-glycerol	Monoglycerides	C23H44O4
163	1-octadecadienoyl-sn-glycerol	Monoglycerides	C21H38O4
164	1-octadecatrienoyl-sn-glycerol	Monoglycerides	C21H36O4
165	1-dodecanoyl-sn-glycerol	Monoglycerides	C15H30O4
166	1-heptadecanoyl-sn-glycerol	Monoglycerides	C20H40O4
167	1-tetradecanoyl-sn-glycerol	Monoglycerides	C17H34O4
168	1-heptadecenoyl-sn-glycerol	Monoglycerides	C20H38O4
169	1-octadecenoyl-sn-glycerol	Monoglycerides	C21H40O4
170	1-hexadecanoyl-sn-glycerol	Monoglycerides	C19H38O4
171	1-octadecanoyl-sn-glycerol	Monoglycerides	C21H42O4
172	Acylated Sterol Glycoside	Others	C53H90O7
173	Ethanol	Others	C2H6O
174	Free Sterol Glycoside	Others	C35H58O6
175	Glycerine	Others	C3H8O3
176	Water	Others	H2O

177	Campesteryl ferulate	Others	C38H56O4
178	Cycloartenyl ferulate	Others	C40H58O4
179	24-Methylene cycloartanyl ferulate	Others	C40H58O4
180	Theobromine-S-D	Others	C7H8N4O2
181	Squalene	Others	C30H50
182	n-hexane	Others	C6H14
183	Methanol	Others	CH4O
184	Acetone	Others	C3H6O
185	1,2,3,6,7,8-Hexachlorooxanthrene	Pesticides	C12H2Cl6O2
186	1,2,3,7,8-Pentachlorooxanthrene	Pesticides	C12H3Cl5O2
187	Dibenzofuran, 2,3,4,7,8-pentachloro-	Pesticides	C12H3Cl5O
188	2,3,7,8-Tetrachlorodibenzo-p-dioxin	Pesticides	C12H4Cl4O2
189	2,3,7,8-Tetrachlorodibenzofuran	Pesticides	C12H4Cl4O
190	benzo(a)pyrene	Pesticides	C20H12
191	Deltamethrin	Pesticides	C22H19Br2NO3
192	Fenitrothion-S-D	Pesticides	C9H12NO5PS
193	2,3,5,6,8,8,9,10,10-nonachlorobornane (parlar 50)	Pesticides	C10H9Cl9
194	1,1'-Biphenyl, 2,3,3',4,4'-pentachloro-	Pesticides	C12H5Cl5
195	1,1'-biphenyl, 2,3',4,4',5-pentachloro-	Pesticides	C12H5Cl5
196	1,1'-Biphenyl, 3,3',4,4',5-pentachloro	Pesticides	C12H5Cl5
197	1,1'-Biphenyl, 2,3,3',4,4',5-hexachloro-	Pesticides	C12H4Cl6
198	Toxaphene	Pesticides	C10H8Cl8
199	1,Oleoyl-sn-glycero-2,Phosphatidylethanolamine	Phospholipids	C23H46NO7P
200	1,Oleoyl-sn-glycero-2,Phosphatidylserine	Phospholipids	C24H44NO9P
201	1,Oleoyl-sn-glycero-2,Phosphatidic acid	Phospholipids	C22H43O7P
202	1,Oleoyl-sn-glycero-2,Phosphatidylcholine	Phospholipids	C26H52NO7P
203	1,Oleoyl-sn-glycero-2,Phosphatidylinositol	Phospholipids	C27H51O13P
204	Phosphatidic acid	Phospholipids	C3H9O6P
205	1,Stearoyl-sn-glycero-2,Phosphatidylethanolamine	Phospholipids	C23H48NO7P
206	1,Stearoyl-sn-glycero-2,Phosphatidylserine	Phospholipids	C24H46NO9P
207	1,Stearoyl-sn-glycero-2,Phosphatidic acid	Phospholipids	C22H45O7P
208	1,Stearoyl-sn-glycero-2,Phosphatidylcholine	Phospholipids	C26H54NO7P
209	1,Stearoyl-sn-glycero-2,Phosphatidylinositol	Phospholipids	C27H53O13P
210	1,Stearoyl-2,Oleoyl-sn-glycero-3,Phosphatidic acid	Phospholipids	C41H79O8P
211	1,Stearoyl-2,Oleoyl-sn-glycero-3,Phosphatidylethanolamine	Phospholipids	C37H73NO4P
212	1,Stearoyl-2,Oleoyl-sn-glycero-3,Phosphatidylserine	Phospholipids	C38H74NO4P
213	Campesterol-octadecenoic acid, ester	Sterol-esters	C46H80O2
214	Campesterol-hexadecanoic acid, ester	Sterol-esters	C44H78O2
215	Sitosterol-octadecenoic acid, ester	Sterol-esters	C47H82O2
216	Sitosterol-hexadecanoic acid, ester	Sterol-esters	C45H80O2
217	Stigmasterol-hexadecanoic acid, ester	Sterol-esters	C45H78O2
218	Stigmasterol-dodecanoic acid, ester	Sterol-esters	C41H70O2
219	Stigmasterol-octadecenoic acid, ester	Sterol-esters	C47H80O2
220	Campesterol	Sterols	C28H48O
221	Cholesterol	Sterols	C27H46O
222	Ergosterol	Sterols	C28H44O
223	Sitosterol	Sterols	C29H50O

224	Stigmasterol	Sterols	C29H48O
225	1,2,3-trioctanoyl-sn-glycerol	Triglycerides	C27H50O6
226	1,2,3-tridecanoyl-sn-glycerol	Triglycerides	C33H62O6
227	1-octanoyl-2-decanoyl-3-dodecanoyl-sn-glycerol	Triglycerides	C33H62O6
228	1-octanoyl-2,3-didodecanoyl-sn-glycerol	Triglycerides	C35H66O6
229	1-decanoyl-2,3-didodecanoyl-sn-glycerol	Triglycerides	C37H70O6
230	1-octanoyl-2-dodecanoyl-3-tetradecanoyl-sn-glycerol	Triglycerides	C37H70O6
231	1,2,3-tridodecanoyl-sn-glycerol	Triglycerides	C39H74O6
232	1-octanoyl-2-dodecanoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C41H76O6
233	1,2-didodecanoyl-3-tetradecanoyl-sn-glycerol	Triglycerides	C41H78O6
234	1-decanoyl-2-dodecanoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C43H80O6
235	1,2-didodecanoyl-3-hexadecanoyl-sn-glycerol	Triglycerides	C43H82O6
236	1,2-didodecanoyl-3-octadecadienoyl-sn-glycerol	Triglycerides	C45H82O6
237	1,2-didodecanoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C45H84O6
238	1-dodecanoyl-2-tetradecanoyl-3-hexadecanoyl-sn-glycerol	Triglycerides	C45H86O6
239	1,2,3-tritetradecanoyl-sn-glycerol	Triglycerides	C45H86O6
240	1-dodecanoyl-2-tetradecanoyl-3-octadecadienoyl-sn-glycerol	Triglycerides	C47H86O6
241	1-dodecanoyl-2-tetradecanoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C47H88O6
242	1-dodecanoyl-2-tetradecanoyl-3-octadecanoyl-sn-glycerol	Triglycerides	C47H90O6
243	1,2-dihexadecanoyl-3-dodecanoyl-sn-glycerol	Triglycerides	C47H90O6
244	1-octadecadienoyl-2,3-ditetradecanoyl-sn-glycerol	Triglycerides	C49H90O6
245	1-dodecanoyl-2-hexadecanoyl-3-octadecadienoyl-sn-glycerol	Triglycerides	C49H90O6
246	1-dodecanoyl-2-hexadecanoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C49H92O6
247	1-hexadecanoyl-2-dodecanoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C49H92O6
248	1-hexadecanoyl-2-octadecenoyl-3-dodecanoyl-sn-glycerol	Triglycerides	C49H92O6
249	1-dodecanoyl-2-hexadecanoyl-3-octadecanoyl-sn-glycerol	Triglycerides	C49H94O6
250	1,2-dioctadecenoyl-3-dodecanoyl-sn-glycerol	Triglycerides	C50H92O6
251	1,2,3-trihexadecenoyl-sn-glycerol	Triglycerides	C51H92O6
252	1-dodecanoyl-2,3-dioctadecenoyl-sn-glycerol	Triglycerides	C51H94O6
253	1-tetradecenoyl-2-hexadecanoyl-3-octadecadienoyl-sn-glycerol	Triglycerides	C51H94O6
254	1-tetradecanoyl-2-hexadecanoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C51H96O6
255	1,2,3-trihexadecanoyl-sn-glycerol	Triglycerides	C51H98O6
256	1,2-dihexadecanoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C53H100O6
257	1,2-dihexadecanoyl-3-octadecanoyl-sn-glycerol	Triglycerides	C53H102O6
258	1-tetradecanoyl-2,3-dioctadecadienoyl-sn-glycerol	Triglycerides	C53H94O6
259	1-tetradecanoyl-2-octadecadienoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C53H96O6
260	1,2-dihexadecenoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C53H96O6
261	1-tetradecanoyl-2,3-dioctadecenoyl-sn-glycerol	Triglycerides	C53H98O6
262	1-hexadecanoyl-2-octadecadienoyl-3-hexadecanoyl-sn-glycerol	Triglycerides	C53H98O6
263	1-hexadecanoyl-2-octadecadienoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C55H100O6
264	1-hexadecanoyl-2-octadecenoyl-3-octadecadienoyl-sn-glycerol	Triglycerides	C55H100O6

265	1-octadecenoyl-2-hexadecanoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C55H102O6
266	1-hexadecanoyl-2,3-dioctadecenoyl-sn-glycerol	Triglycerides	C55H102O6
267	1-hexadecanoyl-2-octadecenoyl-3-octadecanoyl-sn-glycerol	Triglycerides	C55H104O6
268	1-hexadecanoyl-2-octadecanoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C55H104O6
269	1,2-dioctadecadienoyl-3-hexadecenoyl-sn-glycerol	Triglycerides	C55H96O6
270	1-hexadecanoyl-2-octadecatrienoyl-3-octadecadienoyl-sn-glycerol	Triglycerides	C55H96O6
271	1-hexadecanoyl-2,3-dioctadecadienoyl-sn-glycerol	Triglycerides	C55H98O6
272	1-hexadecanoyl-2-octadecatrienoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C55H98O6
273	1-heptadecenoyl-2,3-dioctadecenoyl-sn-glycerol	Triglycerides	C56H102O6
274	1-heptadecanoyl-2,3-dioctadecenoyl-sn-glycerol	Triglycerides	C56H104O6
275	1-octadecenoyl-2,3-dioctadecadienoyl-sn-glycerol	Triglycerides	C57H100O6
276	1,2-dioctadecenoyl-3-octadecatrienoyl-sn-glycerol	Triglycerides	C57H100O6
277	1,2-dioctadecadienoyl-3-octadecanoyl-sn-glycerol	Triglycerides	C57H102O6
278	1-octadecenoyl-2-octadecadienoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C57H102O6
279	1,2,3-trioctadecenoyl-sn-glycerol	Triglycerides	C57H104O6
280	1-octadecanoyl-2-octadecenoyl-3-octadecadienoyl-sn-glycerol	Triglycerides	C57H104O6
281	1-octadecanoyl-2,3-dioctadecenoyl-sn-glycerol	Triglycerides	C57H106O6
282	1-hexadecanoyl-2-octadecenoyl-3-eicosanoyl-sn-glycerol	Triglycerides	C57H108O6
283	1-octadecanoyl-2-octadecenoyl-3-octadecanoyl-sn-glycerol	Triglycerides	C57H108O6
284	1,2,3-trioctadecanoyl-sn-glycerol	Triglycerides	C57H110O6
285	1,2,3-trioctadecatrienoyl-sn-glycerol	Triglycerides	C57H92O6
286	1,2-dioctadecadienoyl-3-octadecatrienoyl-sn-glycerol	Triglycerides	C57H96O6
287	1-octadecenoyl-2,3-dioctadecatrienoyl-sn-glycerol	Triglycerides	C57H96O6
288	1,2,3-trioctadecadienoyl-sn-glycerol	Triglycerides	C57H98O6
289	1-octadecenoyl-2-octadecadienoyl-3-octadecatrienoyl-sn-glycerol	Triglycerides	C57H98O6
290	1-eicosanoyl-2,3-dioctadecadienoyl-sn-glycerol	Triglycerides	C59H106O6
291	1-octadecenoyl-2-octadecadienoyl-3-eicosenoyl-sn-glycerol	Triglycerides	C59H106O6
292	1-eicosanoyl-2-octadecadienoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C59H108O6
293	1-docosanoyl-2-hexadecanoyl-3-octadecadienoyl-sn-glycerol	Triglycerides	C59H108O6
294	1,2-dioctadecenoyl-3-eicosenoyl-sn-glycerol	Triglycerides	C59H108O6
295	1-eicosanoyl-2,3-dioctadecenoyl-sn-glycerol	Triglycerides	C59H110O6
296	1-docosanoyl-2,3-dioctadecadienoyl-sn-glycerol	Triglycerides	C61H110O6
297	1-octadecenoyl-2-octadecadienoyl-3-docosenoyl-sn-glycerol	Triglycerides	C61H110O6
298	1-docosanoyl-2-octadecadienoyl-3-octadecenoyl-sn-glycerol	Triglycerides	C61H112O6
299	1,2-dioctadecenoyl-3-docosenoyl-sn-glycerol	Triglycerides	C61H112O6
300	1-docosanoyl-2,3-dioctadecenoyl-sn-glycerol	Triglycerides	C61H114O6
301	1,2,3-trieicosadienoyl-sn-glycerol	Triglycerides	C63H110O6
302	1,2,3-trieicosenoyl-sn-glycerol	Triglycerides	C63H116O6
303	1,2,3-trieicosanoyl-sn-glycerol	Triglycerides	C63H122O6
304	1,2,3-trieicosapentaenoyl-sn-glycerol	Triglycerides	C63H92O6
305	1,2,3-trieicosatetraenoyl-sn-glycerol	Triglycerides	C63H98O6

306	1,2,3-tridocosenoyl-sn-glycerol	Triglycerides	C69H128O6
307	1,2,3-tridocosanoyl-sn-glycerol	Triglycerides	C69H134O6
308	1,2,3-tritetracosenoyl-sn-glycerol	Triglycerides	C75H140O6
309	1,2,3-tritetracosanoyl-sn-glycerol	Triglycerides	C75H146O6
310	alpha-amyrin	Triterpenealcohols	C30H50O
311	beta-amyrin	Triterpenealcohols	C30H50O
312	Butyrospermol	Triterpenealcohols	C30H50O
313	Dihydrolupeol	Triterpenealcohols	C30H52O
314	Germanicol	Triterpenealcohols	C30H50O
315	Lupeol	Triterpenealcohols	C30H50O
316	4-ethylphenol	Triterpenealcohols	C8H10O
317	Tyrosol	Triterpenealcohols	C8H10O2
318	Ubiquinone q6	Ubiquinones	C39H58O4
319	Ubiquinone q7	Ubiquinones	C44H66O4
320	Ubiquinone q8	Ubiquinones	C49H74O4
321	Ubiquinone q9	Ubiquinones	C54H82O4
322	Ubiquinone q10	Ubiquinones	C59H90O4
323	Alpha-Tocopherol	Vitamin E	C29H50O2
324	Alpha-Tocotrienol	Vitamin E	C29H44O2
325	Beta-Tocopherol	Vitamin E	C28H48O2
326	Beta-Tocotrienol	Vitamin E	C28H42O2
327	Gamma-Tocopherol	Vitamin E	C27H46O2
328	Gamma-Tocotrienol	Vitamin E	C27H40O2
329	Delta-Tocopherol	Vitamin E	C28H48O2
330	Delta-Tocotrienol	Vitamin E	C28H42O2
331	Plastochromanol-8	Vitamin E	C52H80O2

Table A.2 Experimental data of primary properties

Compound	Property						
	Melting point	Boiling point	Critical temperature	Critical pressure	Critical volume	Gibbs energy of formation	Enthalpy of formation
Fatty acids	18	14	13	13	8	8	13
Triglycerides							
Diglycerides							
Monoglycerides	1	1					
Fatty esters	30	19	2	1	1	1	7
Others							
Fatty alcohols	12	9	7	7	6		7

Table A.3 Experimental data of temperature dependent properties

Compound	Property					
	Vapor pressure	Liquid heat capacity	Liquid density	Liquid viscosity	Surface tension	Liquid thermal conductivity
Fatty acids	20	17	11	12	9	8
Triglycerides	7	12		16	9	
Diglycerides					1	
Monoglycerides	6				5	
Fatty esters	14	15	26	21	17	
Others						
Fatty alcohols		1				

A.4 Consistency test of pure compound properties

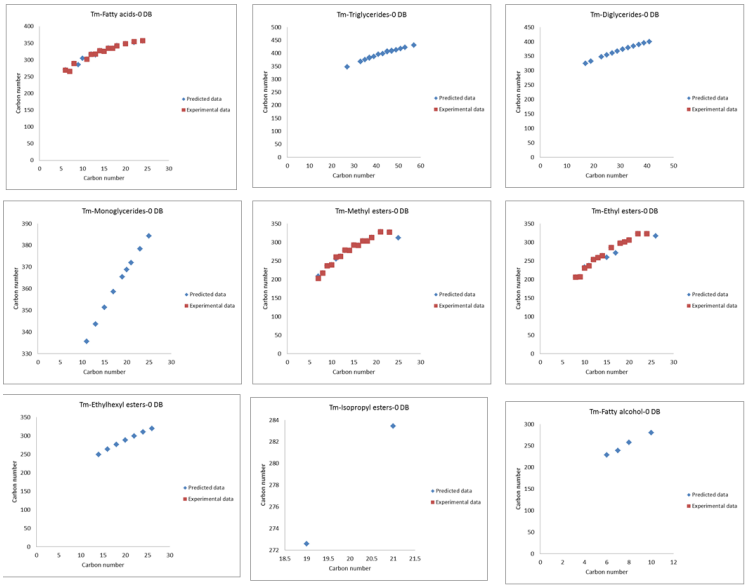


Figure A.4.1 Consistency of the melting point property model

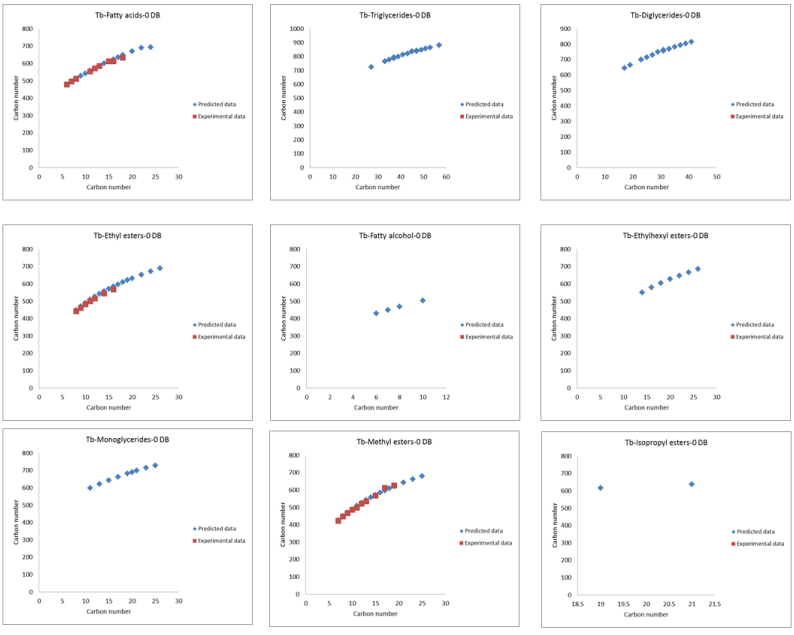


Figure A.4.2 Consistency of the boiling point property model

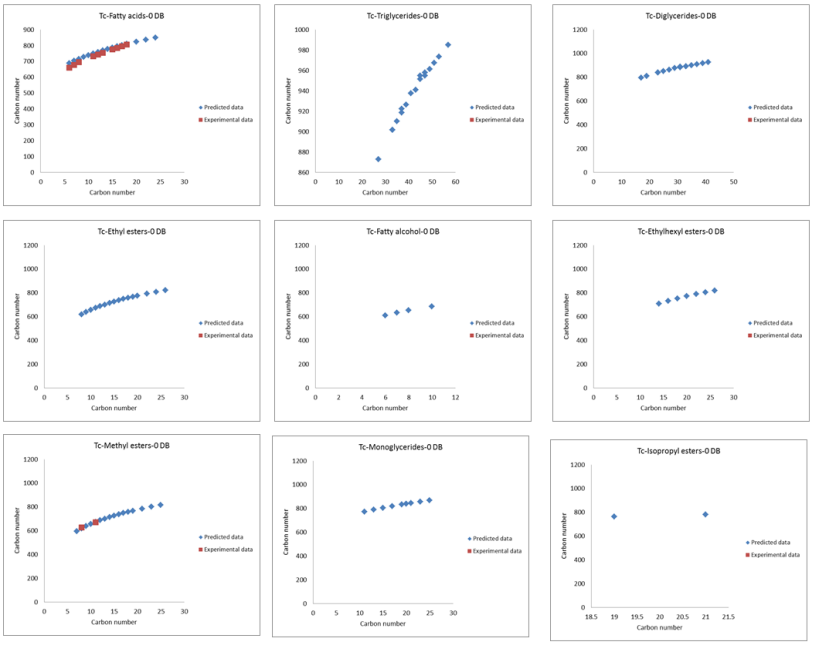


Figure A.4.3 Consistency of the critical temperature property model

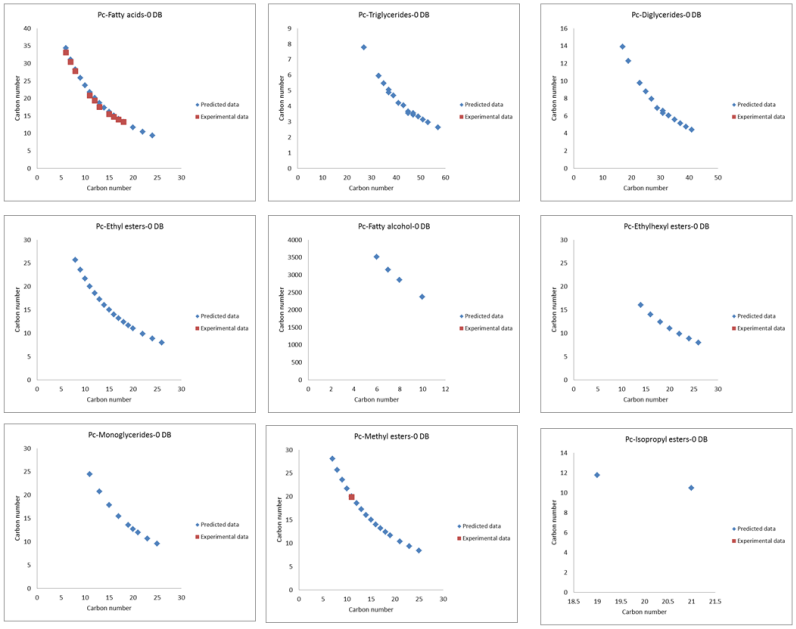


Figure A.4.4 Consistency of the critical pressure property model

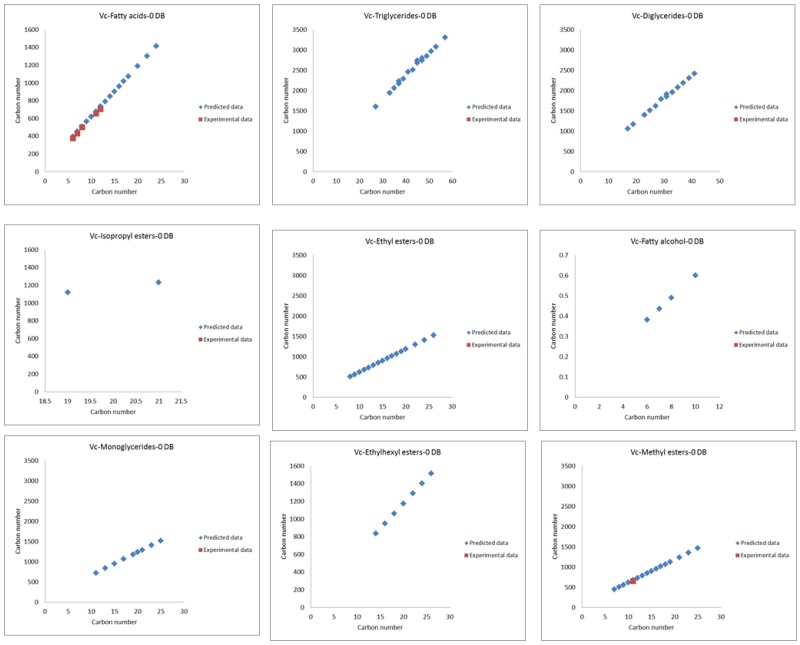


Figure A.4.5 Consistency of the critical volume property model

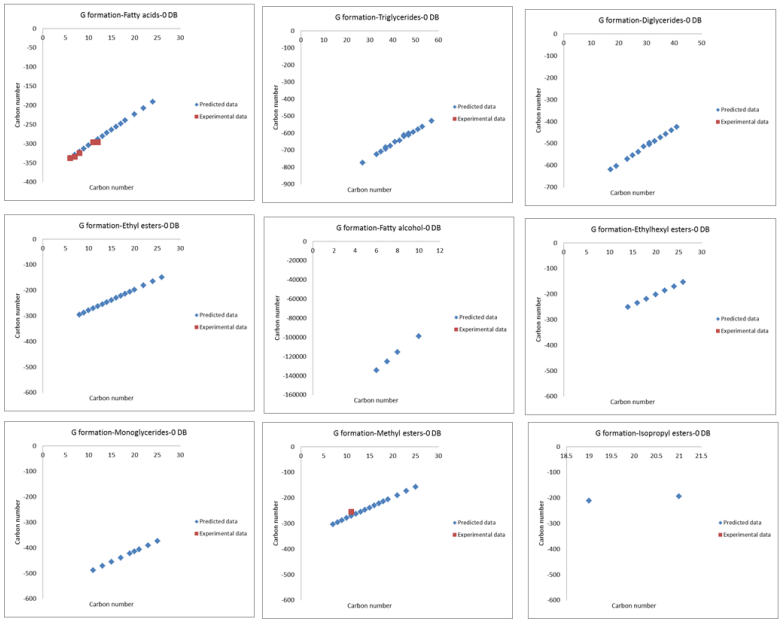


Figure A.4.6. Consistency of the ΔG formation at 298 K property model

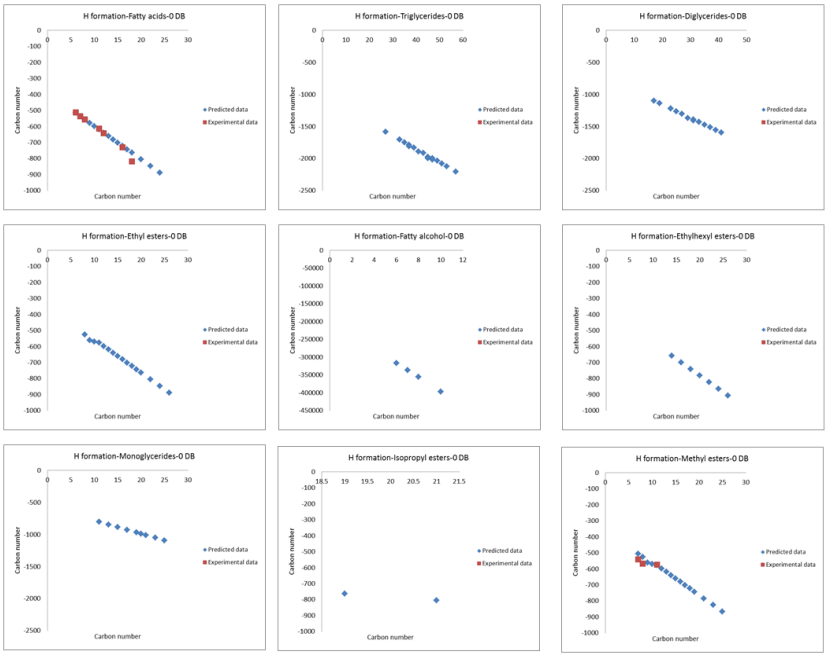


Figure A.4.7. Consistency of the ΔH formation at 298 K property model

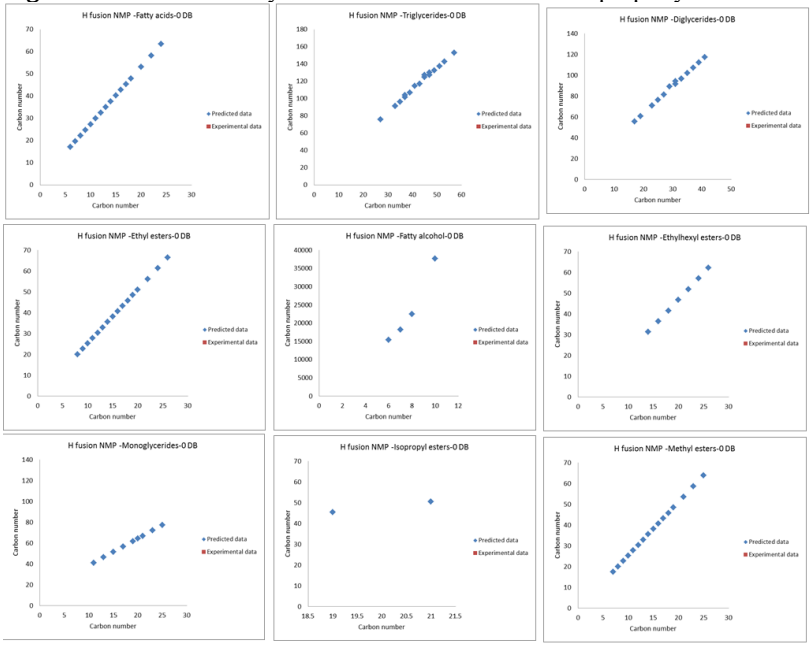


Figure A.4.8. Consistency of the ΔH fusion at 298 K property model

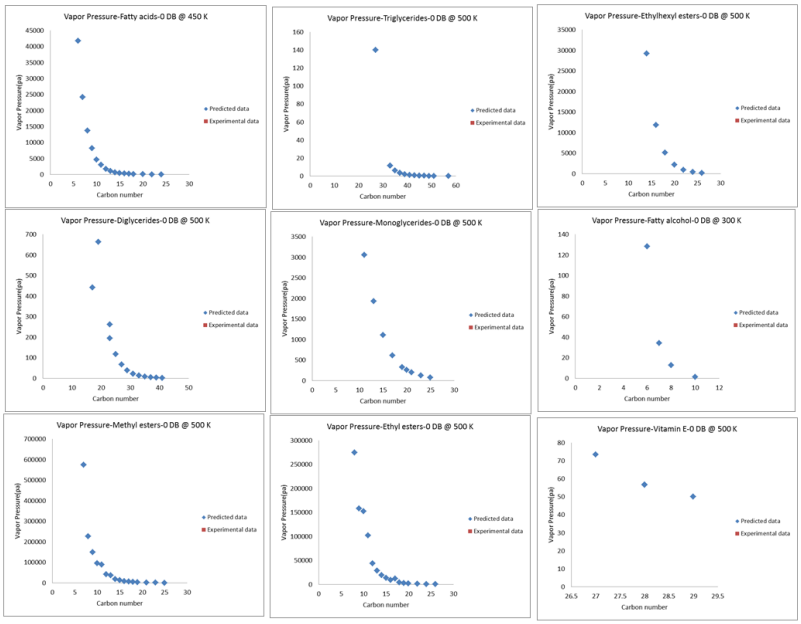


Figure A.4.9. Consistency of vapor pressure property models

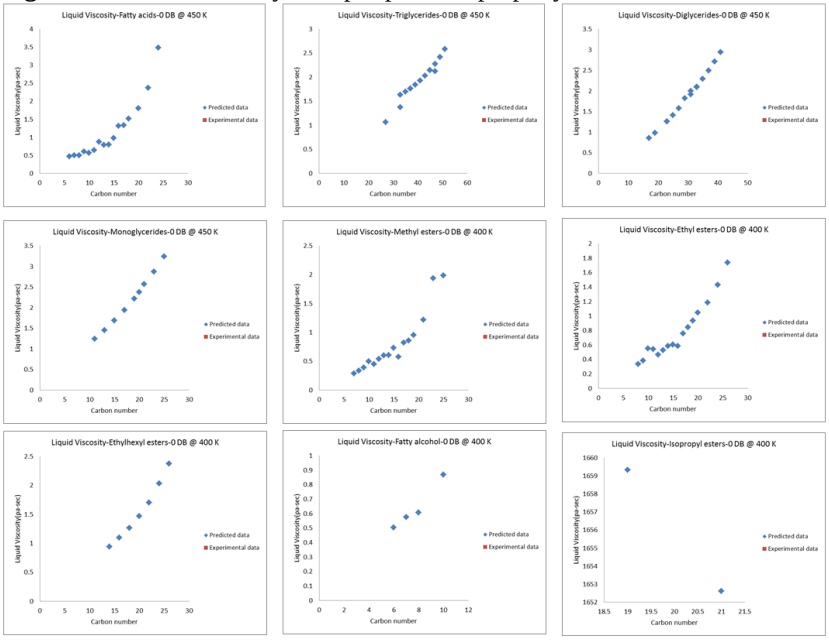


Figure A.4.10. Consistency of liquid viscosity property models

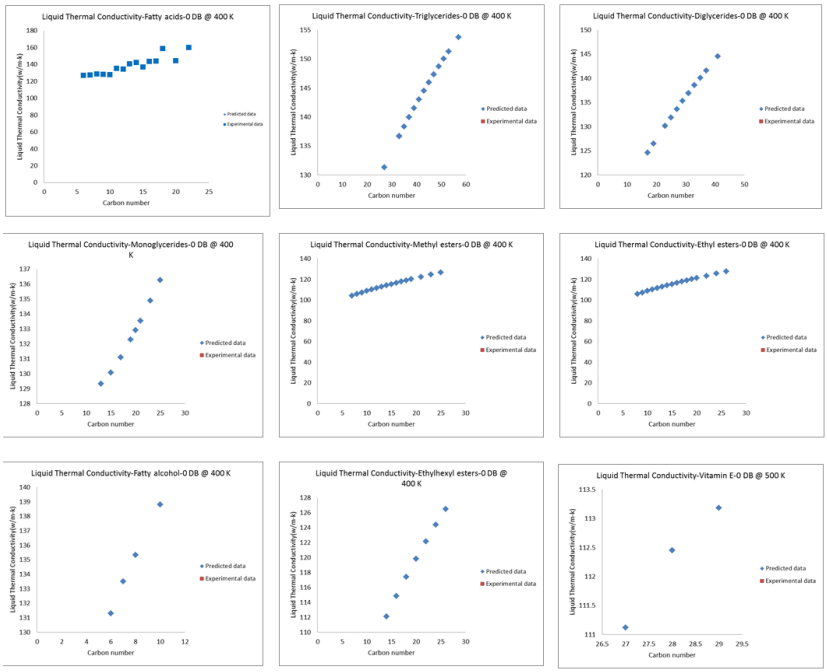


Figure A.4.11. Consistency of liquid thermal conductivity models

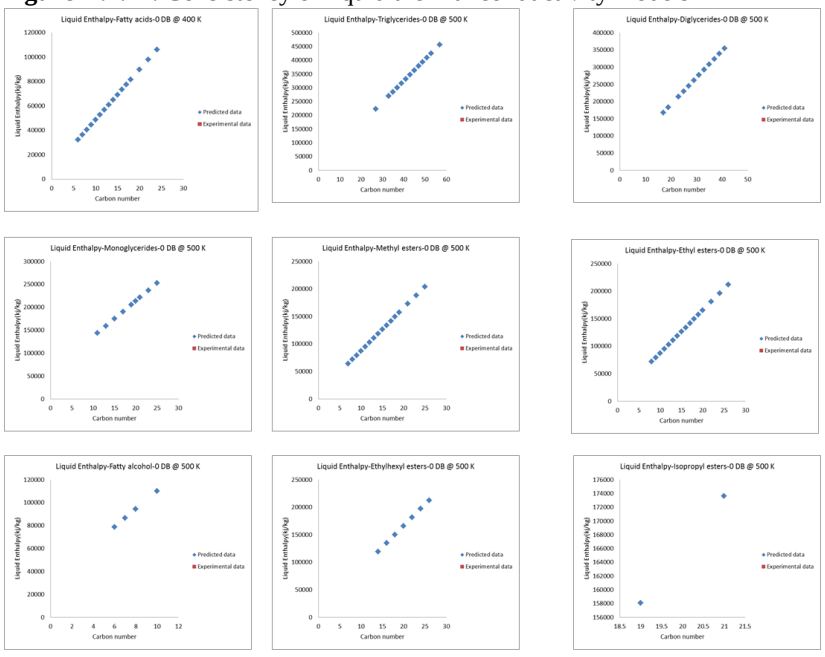


Figure A.4.12. Consistency of liquid enthalpy property models

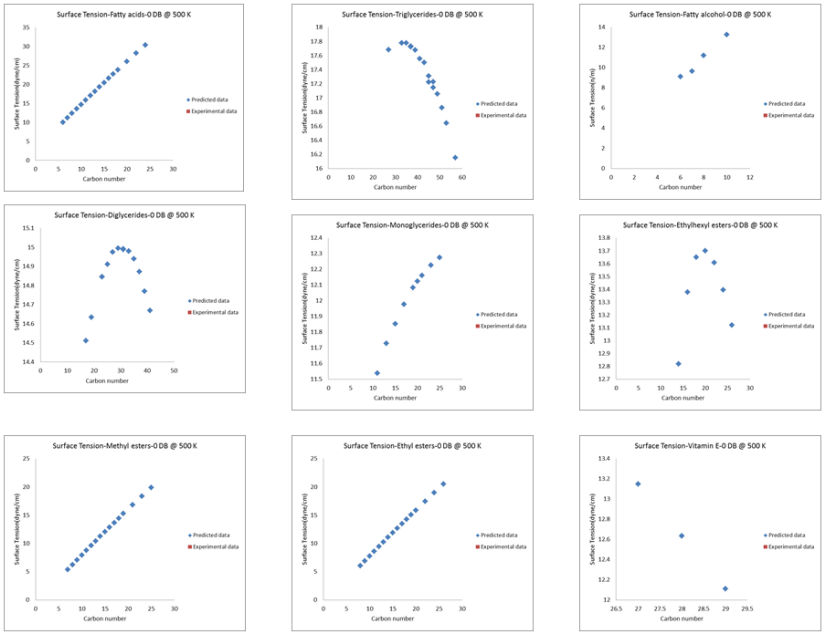


Figure A.4.13. Consistency of surface tension property models

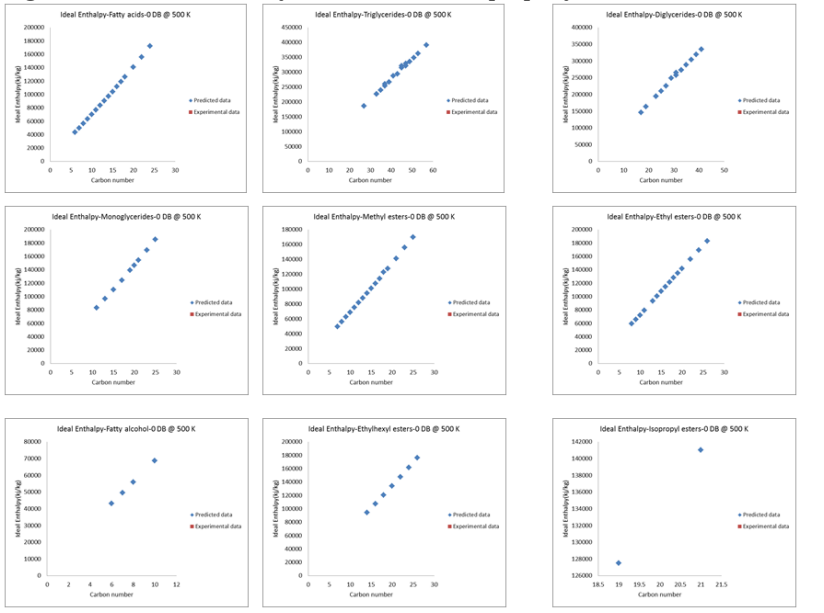


Figure A.4.14. Consistency of the ideal enthalpy property model

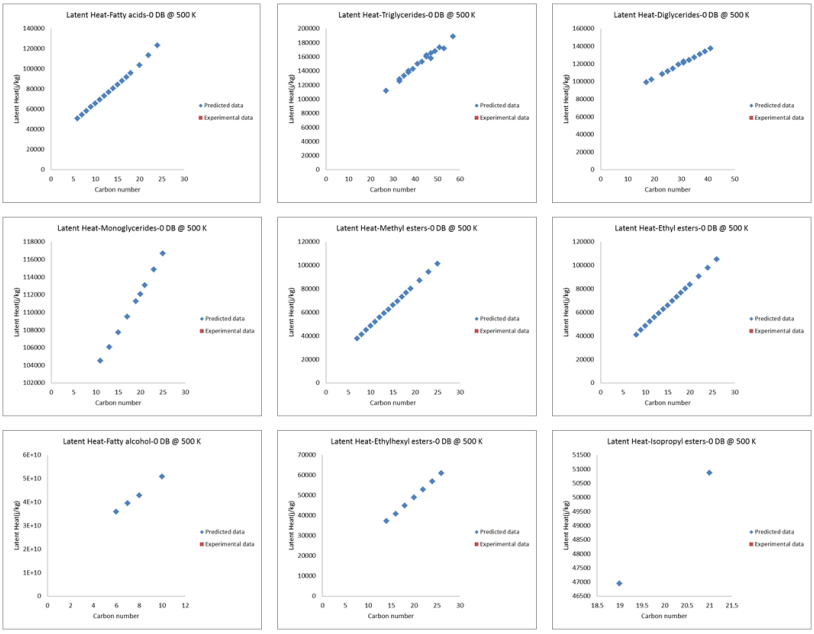


Figure A.4.15. Consistency of the latent heat property model

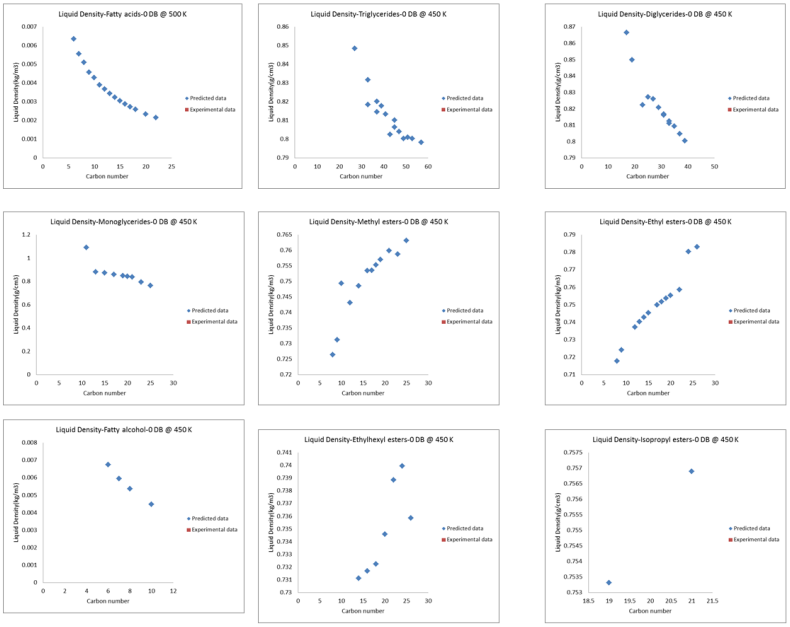


Figure A.4.16. Consistency of liquid density models

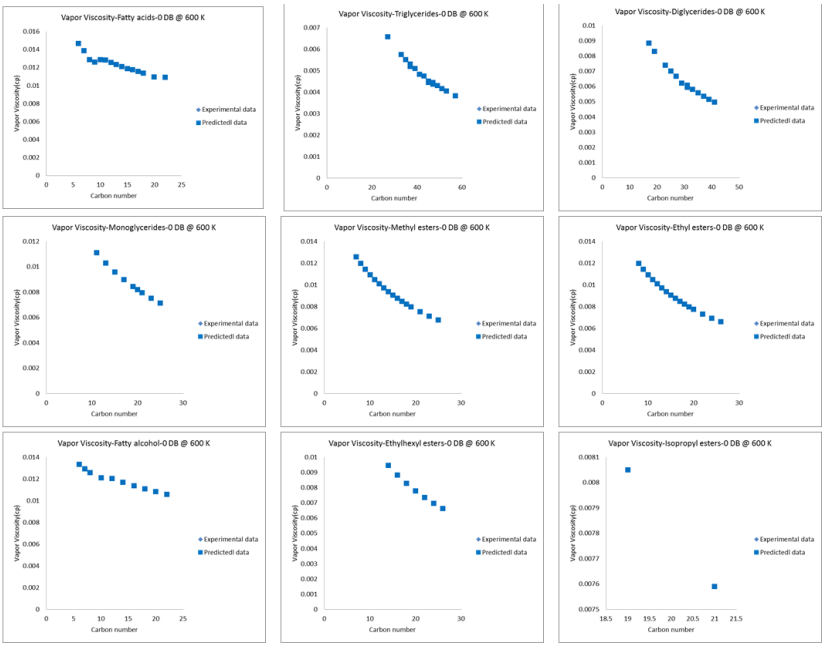


Figure A.4.17. Consistency of the vapor viscosity model

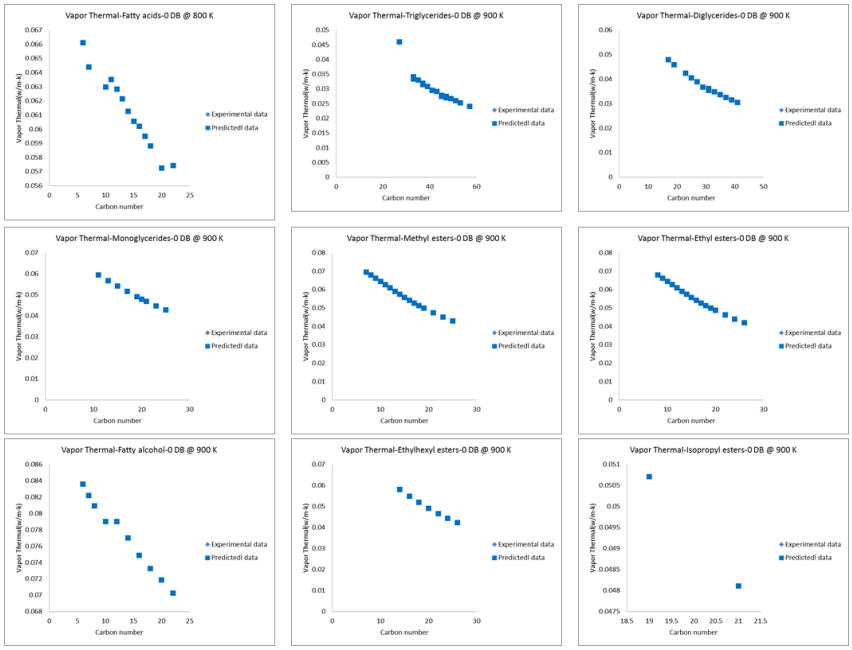


Figure A.4.18. Consistency of the vapor thermal conductivity model

APPENDIX B

Results from database search

Table B.1 List of solvent candidates in problem 5.3 through database search

Id	Chemname	Casno	Mw	Tb	SolPar
1	3-Pentanol,3-methyl-	000077-74-7	102.18	395.55	17.5
2	Naphthalene,decahydro-	000091-17-8	138.25	428.65	18
3	1,2-Ethanediamine,N,N-diethyl-	000100-36-7	116.21	417.15	17.5
4	Morpholine,4-ethyl-	000100-74-3	115.18	411.65	18.3
5	2-Pentanone,4-methoxy-4-methyl-	000107-70-0	130.19	433.15	17.2
6	1-Butanol,3-methyl-,formate	000110-45-2	116.16	396.65	16.4
7	1-Butanamine,N-methyl-	000110-68-9	87.16	364.15	17
8	Ethane,1,1',1"--methylidynetris(oxy)-tris-	000122-51-0	148.2	416.15	17
9	Pentanal,2-methyl-	000123-15-9	100.16	390.15	17.7
10	Butane,1-iodo-	000542-69-8	184.02	403.75	17.6
11	Hexane,1-chloro-	000544-10-5	120.62	408.15	17.2
12	Butanal,3-methyl-	000590-86-3	86.13	365.65	17.7
13	Furan,2,5-dimethyl-	000625-86-5	96.13	366.65	17.9
14	Pentane,1-iodo-	000628-17-1	198.05	428.15	17.2
15	Cyclohexanamine,N-ethyl-	005459-93-8	127.23	437.15	17.6
16	1-Butanol,2-ethyl-,acetate	010031-87-5	144.21	435.65	17
17	2-Butenoic-acid,ethyl-ester	010544-63-5	114.14	409.65	18.2
18	1-HEXANAL	000066-25-1	100.16	404.15	18.1497
19	ISOPROPYL-IODIDE	000075-30-9	169.99	362.65	17.7838
20	BROMOTRICHLOROMETHANE	000075-62-7	198.27	378.15	18.32
21	3,3-DIMETHYL-2-BUTANONE	000075-97-8	100.16	379.25	16.9229
22	1,1,1,2-TETRACHLORODIFLUOROETHANE	000076-11-9	203.83	364.65	16.5142
23	1,1,2,2-TETRACHLORODIFLUOROETHANE	000076-12-0	203.83	366.15	16.42
24	1,1-DICHLOROPROPANE	000078-99-9	112.99	361.25	18.1222
25	ETHYL-ISOBUTYRATE	000097-62-1	116.6	383.25	16.504
26	ETHYL-METHACRYLATE	000097-63-2	114.14	390.15	17.3537
27	ISOBUTYL-METHACRYLATE	000097-86-9	142.2	428.15	16.5772

28	n-BUTYL-METHACRYLATE	000097-88-1	142.2	433.15	17.1406
29	BENZOTRIFLUORIDE	000098-08-8	146.11	375.25	16.8796
30	p-CHLOROBENZOTRIFLUORIDE	000098-56-6	180.56	411.65	16.7135
31	CUMENE	000098-82-8	120.19	425.55	17.4362
32	alpha-METHYLSTYRENE	000098-83-9	118.18	438.55	18.3262
33	VINYLCYCLOHEXENE	000100-40-3	108.18	401.15	16.9054
34	ETHYLBENZENE	000100-41-4	106.17	409.25	17.9848
35	n-PROPYLBENZENE	000103-65-1	120.19	432.35	17.6612
36	ETHYL-PROPIONATE	000105-37-3	102.13	372.25	17.736
37	VINYL-PROPIONATE	000105-38-4	100.12	364.35	18.1159
38	sec-BUTYL-ACETATE	000105-46-4	116.16	385.15	16.7816
39	ETHYL-n-BUTYRATE	000105-54-4	116.16	394.65	17.3779
40	n-PROPYL-n-BUTYRATE	000105-66-8	130.19	416.15	16.8367
41	3-HEPTANONE	000106-35-4	114.19	420.15	18.1076
42	n-PROPYL-PROPIONATE	000106-36-5	116.16	395.65	17.5677
43	p-XYLENE	000106-42-3	106.17	411.45	17.9031
44	ISOBUTYL-ACRYLATE	000106-63-8	128.17	405.15	17.1219
45	n-PROPYL-IODIDE	000107-08-4	169.99	375.75	18.3288
46	2-PENTANONE	000107-87-9	86.13	375.35	18.2948
47	METHYL-ISOBUTYL-KETONE	000108-10-1	100.16	389.65	17.4328
48	ISOPROPYL-ACETATE	000108-21-4	102.13	361.75	17.1536
49	m-XYLENE	000108-38-3	106.17	412.25	18.0535
50	ETHYL-ISOVALERATE	000108-64-5	130.19	408.15	16.5203
51	MESITYLENE	000108-67-8	120.19	437.85	17.9647
52	TOLUENE	000108-88-3	92.14	383.75	18.3242
53	n-BUTYL-n-BUTYRATE	000109-21-7	144.21	439.15	16.973
54	n-PROPYL-ACETATE	000109-60-4	102.13	374.65	17.8885
55	n-BUTYL-MERCAPTAN	000109-79-5	90.19	371.65	17.8002
56	5-METHYL-2-HEXANONE	000110-12-3	114.19	417.15	17.7375
57	ISOBUTYL-ACETATE	000110-19-0	116.16	389.65	17.0469
58	2-HEPTANONE	000110-43-0	114.19	424.15	17.8842
59	n-PENTYLAMINE	000110-58-7	87.16	377.45	17.89
60	n-PENTYL-MERCAPTAN	000110-66-7	104.22	399.75	17.654
61	PIPERAZINE	000110-85-0	86.14	419.15	17.962
62	n-HEXYLAMINE	000111-26-2	101.19	405.95	17.6285
63	n-HEXYL-MERCAPTAN	000111-31-9	118.24	424.15	17.4482
64	BUTYL-VINYL-ETHER	000111-34-2	100.16	367.15	16.4617
65	DI-n-PROPYL-SULFIDE	000111-47-7	118.24	416.05	17.1108
66	n-HEPTYLAMINE	000111-68-2	115.22	429.15	17.3321
67	1-HEPTANAL	000111-71-7	114.19	425.95	17.9533
68	DI-n-BUTYLAMINE	000111-92-2	129.25	432.75	16.5967
69	2-ETHYLHEXANAL	000123-05-7	128.21	436.15	17.561
70	4-HEPTANONE	000123-19-3	114.19	417.15	17.2289
71	PARALDEHYDE	000123-63-7	132.16	397.45	17.7283

72	n-BUTYL-ACETATE	000123-86-4	116.16	399.25	17.5858
73	ISOPENTYL-ACETATE	000123-92-2	130.19	415.65	16.9317
74	DIALLYLAMINE	000124-02-7	97.16	384.15	17.8703
75	beta-PINENE	000127-91-3	136.24	439.15	16.7839
76	ETHYL-ACRYLATE	000140-88-5	100.12	372.55	18.1815
77	n-BUTYL-ACRYLATE	000141-32-2	128.17	418.15	17.9164
78	DI-n-PROPYLAMINE	000142-84-7	101.19	382.45	16.7019
79	CYCLOHEPTANE	000291-64-5	98.19	391.55	17.2137
80	CYCLOOCTANE	000292-64-8	112.22	422.15	17.3733
81	DIETHYL-SULFIDE	000352-93-2	90.19	365.25	17.5074
82	ISOBUTYL-MERCAPTAN	000513-44-0	90.19	361.65	17.2439
83	ISOBUTYL-FORMATE	000542-55-2	102.13	371.35	17.8235
84	1-CHLOROPENTANE	000543-59-9	106.6	380.95	16.9524
85	METHYL-ISOBUTYRATE	000547-63-7	102.13	365.65	17.4714
86	METHYL-ISOPROPYL-KETONE	000563-80-4	86.13	367.45	18.2533
87	ETHYL-ISOPROPYL-KETONE	000565-69-5	100.16	386.65	17.3251
88	DIISOPROPYL-KETONE	000565-80-0	114.19	398.55	17.8807
89	3-HEXANONE	000589-38-8	100.16	396.65	17.8979
90	sec-BUTYL-FORMATE	000589-40-2	102.13	370.15	17.3165
91	n-BUTYL-PROPIONATE	000590-01-2	130.19	419.95	17.3769
92	2-HEXANONE	000591-78-6	100.16	400.75	18.1374
93	ALLYL-ACETATE	000591-87-7	100.12	376.65	17.9481
94	n-BUTYL-FORMATE	000592-84-7	102.13	379.25	18.2255
95	o-ETHYLTOLUENE	000611-14-3	120.19	438.35	18.0132
96	1,2-DICHLOROBUTANE	000616-21-7	127.01	397.25	18.1591
97	m-ETHYLTOLUENE	000620-14-4	120.19	434.45	17.8567
98	p-ETHYLTOLUENE	000622-96-8	120.19	435.15	17.783
99	METHYL-n-BUTYRATE	000623-42-7	102.13	375.95	18.0264
100	METHYL-n-BUTYL-SULFIDE	000628-29-5	104.22	396.65	17.5661
101	n-PENTYL-ACETATE	000628-63-7	130.19	422.35	17.6089
102	CYCLOHEPTENE	000628-92-2	96.17	388.15	17.346
103	1,2-DIETHOXYETHANE	000629-14-1	118.18	392.55	17.0045
104	n-HEXYL-FORMATE	000629-33-4	130.19	428.65	17.7755
105	n-PENTYL-FORMATE	000638-49-3	116.16	403.55	17.9777
106	n-PROPYL-ISOBUTYRATE	000644-49-5	130.19	408.55	16.5263
107	2,5-DIMETHYL-2,4-HEXADIENE	000764-13-6	110.2	407.65	16.7151
108	n-PROPYLCYCLOHEXANE	001678-92-8	126.24	429.85	16.348
109	n-BUTYLCYCLOPENTANE	002040-95-1	126.24	429.75	16.3854
110	n-PROPYL-METHACRYLATE	002210-28-8	128.17	414.15	17.1902
111	METHYL-n-PROPYL-SULFIDE	003877-15-4	90.19	368.75	17.7109
112	3-METHYLHEXANAL	019269-28-4	114.19	416.15	17.2812

Table B.2 List of additives for jet-fuel blend problem **5.3.1**

No.	Compound	Formula
1	n-Butanol	C4H10O
2	n-Hexanol	C6H14O
3	Butyl butyrate	C8H16O2
4	Ethyl octanoate	C10H20O2
5	Ethyl cyclohexane	C8H16
6	Limonene	C10H16
7	Decane	C10H22
8	Nonane	C9H20
9	n-OCTANE	C8H18
10	n-HEPTANE	C7H16
11	n-Hexane	C6H14
12	Dodecane	C12H26
13	Tetradecane	C14H30
14	Hexadecane	C16H34
15	Octadecane	C18H38
16	EIcosane	C20H42

Table B.3 List of additives for jet-fuel blend problem **5.3.2**

No.	Compound	Formula
1	2-ethyl-1-Hexanol	C8H18O
2	2-OCTANOL	C8H18O
3	2,2-dimethyl-3-Pentanol	C7H16O
4	Cyclohexane, 1,1'-(1,2-ethanediyl)bis-	C14H26
5	6-Methyl-2-heptanol	C8H18O
6	4-Methyl-3-heptanol	C8H18O
7	3-Heptanol, 5-methyl-	C8H18O
8	2-Nonanone	$C_9H_{18}O$
9	Triethylene glycol	C6H14O4
10	2-heptanol	C7H16O
11	2-Butanol, 3,3-dimethyl-	C6H14O
12	3,3,4,4-tetramethylhexane	C10H22
13	3-methyl-2-hexanol	C7H16O
14	3,4-dimethyl-2-pentanol	C7H16O
15	Cyclopentane, decyl-	C15H30
16	2,2-DIMETHYL-1-PROPANOL	C5H12O
17	2,2,3,3-tetramethylhexane	C10H22
18	Cyclohexane, octyl-	C14H28
19	3-methylpentadecane	C16H34
20	tert-Butylcyclohexane	C10H20

21	Pentadecane	C15H32
22	2-METHYL-1-PENTANOL	C6H14O
23	2-Ethyl-1-butanol	C6H14O
24	Ethanedioic acid, diethyl ester	C6H10O4
25	1-Pentanol	C5H12O
26	3,3,6,6-Tetramethyloctane	C12H26
27	Propanedioic acid, diethyl ester	C7H12O4
28	n-Nonylcyclopentane	C14H28
29	Cyclodecane	C10H20
30	METHYL-ACETOACETATE	C5H8O3
31	2,2,3,3-tetramethylpentane	C9H20
32	4-METHYL-2-PENTANOL	C6H14O
33	Heptylcyclohexane	C13H26
34	1-Methylcyclopropanemethanol	C5H10O
35	2-HEPTANONE	C7H14O
36	3-METHYL-1-BUTANOL	C5H12O
37	3-METHYLHEXANAL	C7H14O
38	Tetradecane	C14H30
39	2-METHYL-1-BUTANOL	C5H12O
40	2-PENTANOL	C5H12O
41	3-PENTANOL	C5H12O
42	3-METHYL-2-BUTANOL	C5H12O
43	Carbonic acid, diethyl ester	C5H10O3
44	2,2,4,4,6-Pentamethylheptane	C12H26
45	N-OCTYLCYCLOPENTANE	C13H26
46	PROPYLENE-GLYCOL-MONOMETHYL-ETHER-ACETATE	$C_6H_{12}O_3$
47	n-BUTYL-METHACRYLATE	C8H14O2
48	Cyclopentane, (1,1-dimethylethyl)-	C9H18
49	Cyclohexane, hexyl-	C12H24
50	Tridecane	C13H28
51	2,2,4,6,6-Pentamethylheptane	C12H26
52	Cyclopentane, 1,1,3,3-tetramethyl-	C9H18
53	Spiro[2.5]octane	C8H14
54	ISOBUTYL-METHACRYLATE	C8H14O2
55	n-Heptylcyclopentane	C12H24
56	1,3-Dimethylbicyclo[1.1.0]butane	C6H10
57	n-PROPYL-METHACRYLATE	C7H12O2
58	Cyclohexane, 1,1,3-trimethyl-	C9H18
59	Cyclohexane, 1,1,2-trimethyl-	C9H18
60	Dodecane	C12H26
61	2,2,5,5-tetramethylhexane	C10H22
62	2-methylundecane	C12H26
63	Cyclohexane, (1-methylethyl)-	C9H18

64	Cyclohexane, butyl-	C10H20
65	4-methylundecane	C12H26
66	Toluene	
67	2,2,3,4,4-PENTAMETHYLPENTANE	C10H22
68	1,1-Dimethylcyclohexane	C8H16
69	Cyclooctane	C8H16
70	Undecane	C11H24
71	Benzene, propyl-	C9H12
72	2,2-dimethyloctane	C10H22
73	2,2,4,4-tetramethylpentane	C9H20
74	3,3,5-trimethylheptane	C10H22
75	Cyclopentane, pentyl-	C10H20
76	1,1'-Bicyclopropyl	C6H10
77	4,5-diethyloctane	C12H26
78	2-Propenoic acid, 2-methyl-, methyl ester	C5H8O2
79	2-methyldecane	C11H24
80	1,1,2-Trimethylcyclopentane	C8H16
81	Cyclopentane, 1,1,3-trimethyl-	C8H16
82	3,3-dimethylheptane	C9H20
83	3,3-diethylpentane	C9H20
84	4,4-dimethylheptane	C9H20
85	3-methyl-3-ethylhexane	C9H20
86	Cyclohexane, propyl-	C9H18
87	METHYLCYCLOHEPTANE	C8H16
88	4-Methyldecane	C11H24
89	2,2,6-Trimethylheptane	C10H22
90	Butanoic acid, propyl ester	C7H14O2
91	Decane	C10H22
92	Ethylbenzene	C8H10
93	2,2-dimethylheptane	C9H20
94	2,3,3-trimethylhexane	C9H20
95	1,2-DIMETHOXYETHANE	C4H10O2
96	2,4,4-trimethylhexane	C9H20
97	3,3,4-trimethylhexane	C9H20
98	1-METHYLETHYL-CYCLOPENTANE	C8H16
99	Furan, tetrahydro-2-methyl-	C5H10O
100	Furan, tetrahydro-3-methyl-	C5H10O
101	Cyclopentane, butyl-	C9H18
102	1,1,2,2-tetramethylcyclopropane	C7H14
103	n-Butyl ether	C8H18O
104	1,1-DIMETHYLCYCLOPENTANE	C7H14
105	TRANS-1,2-Dimethylcyclohexane	C8H16
106	CIS-1,2-Dimethylcyclohexane	C8H16
107	cis-1,3-dimethylcyclohexane	C8H16

108	CIS-1,4-DIMETHYLCYCLOHEXANE	C8H16
109	trans-1,3-Dimethylcyclohexane	C8H16
110	TRANS-1,4-DIMETHYLCYCLOHEXANE	C8H16
111	Cyclohexane, ethenyl-	C_8H_{14}
112	CYCLOHEPTANE	C7H14
113	2,3,3,4-tetramethylpentane	C9H20
114	3-METHYL-3-ETHYLPENTANE	C8H18
115	Cyclohexane, ethylidene-	C_8H_{14}
116	Ethylcyclohexane	C8H16
117	3-Methylnonane(DL)	C10H22
118	3-ethyloctane	C10H22
119	4-Methylnonane(DL)	C10H22
120	4-ethyloctane	C10H22
121	4-propylheptane	C10H22
122	5-methylnonane	C10H22
123	2,2,5-trimethylhexane	C9H20
124	2,2,3-trimethylhexane	C9H20
125	2,3-dimethyloctane	C10H22
126	2,2,3,4-tetramethylpentane	C9H20
127	Butanoic acid, ethyl ester	C6H12O2
128	2-Cyclopropylhexane	C9H18
129	Ethane, 1,2-diethoxy-	C6H14O2
130	1,2-DIETHOXYETHANE	C6H14O2
131	2,2,4-trimethylhexane	C9H20
132	2,2-dimethyl-3-ethylpentane	C9H20
133	2-ETHOXYETHANOL	C4H10O2
134	Nonane	C9H20
135	Cyclopentane, (2-methylpropyl)-	C9H18
136	1-ETHYL-1-METHYLCYCLOPENTANE	C8H16
137	3,4,5-TRIMETHYLHEPTANE	C10H22
138	2,2-DIMETHYLHEXANE	C8H18
139	Propane, 2-methyl-2-(1-methylethoxy)-	C7H16O
140	2,7-dimethyloctane	C10H22
141	METHYL-tert-BUTYL-ETHER	C5H12O
142	2,4-Dimethyloctane	C10H22
143	2,5-dimethyloctane	C10H22
144	2,6-Dimethyloctane	C10H22
145	Propylcyclopentane	C8H16
146	Methylcyclohexane	C7H14
147	3,4-DIETHYLHEXANE	C10H22
148	1-Methoxy-2-propanol	C4H10O2
149	2-methyloctane	C9H20
150	Ethane, 1,1-diethoxy-	C6H14O2
151	3,3-DIMETHYLPENTANE	C7H16

152	3-methyloctane	C9H20
153	3-ethylheptane	C9H20
154	4-methyloctane	C9H20
155	4-ethylheptane	C9H20
156	2,2,4-trimethylpentane	C8H18
157	2,2,3-TRIMETHYLPENTANE	C8H18
158	2,3-dimethylheptane	C9H20
159	ETHYLAL	C5H12O2
160	2,2'-oxybisbutane	C8H18O
161	3,4-DIMETHYLHEPTANE	C9H20
162	n-OCTANE	C8H18
163	Isopropylcyclobutane	C7H14
164	2,2-dimethylpentane	C7H16
165	2,6-dimethylheptane	C9H20
166	trans-1,2-Dimethylcyclopentane	C7H14
167	1,2-Dimethylcyclopentane (cis-)	C7H14
168	trans-1,3-DIMETHYLCYCLOPENTANE	C7H14
169	cis-1,3-DIMETHYLCYCLOPENTANE	C7H14
170	CYCLOHEXANE	C6H12
171	2,4-dimethylheptane	C9H20
172	2,5-dimethylheptane	C9H20
173	2-methyl-3-ethylhexane	C9H20
174	2-methyl-4-ethylhexane	C9H20
175	2,3,5-trimethylhexane	C9H20
176	3,5-DIMETHYLHEPTANE	C9H20
177	3-ethyl-4-methylhexane	C9H20
178	2-methylheptane	C8H18
179	2,2,3-trimethylbutane	C7H16
180	METHYL-ISOBUTYL-ETHER	C5H12O
181	3-METHYLHEPTANE	C8H18
182	3-ethylhexane	C8H18
183	2,3-DIMETHYLHEXANE	C8H18
184	2,3,4-TRIMETHYLPENTANE	C8H18
185	METHYL-sec-BUTYL-ETHER	C5H12O
186	3,4-DIMETHYLHEXANE	C8H18
187	2,4-dimethyl-3-ethylpentane	C9H20
188	1,1,2-TRIMETHYLCYCLOPROPANE	C6H12
189	Methylcyclopentane	C6H12
190	2,2-dimethylbutane	C6H14
191	2,5-DIMETHYLHEXANE	C8H18
192	2,4-DIMETHYLHEXANE	C8H18
193	2-methyl-3-ethylpentane	C8H18
194	DIISOPROPYL-ETHER	C6H14O
195	Isopropylcyclopropane	C6H12

196	3-METHYLHEXANE	C7H16
197	3-ethylpentane	C7H16
198	Butylcyclopropane	C7H14
199	Ethylcyclobutane	C6H12
200	2,4-DIMETHYLPENTANE	C7H16
201	1-ETHYL-1-METHYLCYCLOPROPANE	C6H12
202	2-METHYLPENTANE	C6H14
203	1,1'-methylenebiscyclopropane	C7H12
204	2,3-DIMETHYLBUTANE	C6H14
205	3-METHYLPENTANE	C6H14
206	1,2,3-trimethylcyclopropane	C6H12
207	Cyclopropane, (1-methylpropenyl)-	C_7H_{12}
208	1,1-diethylcyclopropane	C7H14
209	Cyclopropane, (1-ethylvinyl)-	C_7H_{12}

Table B.4 List of additives for diesel blend problem **5.4**

No.	Compounds	Formula
1	1,2-BUTANEDIOL	C4H10O2
2	1,2-PROPYLENE-GLYCOL	C3H8O2
3	1,4-Butanediol	C4H10O2
4	1-BUTANOL	C4H10O
5	1-hexanol	C6H14O
6	1-Pentanol	C5H12O
7	2-(2-METHOXYETHOXY)ETHANOL	C5H12O3
8	2,3-BUTANEDIOL	C4H10O2
9	2-BUTOXYETHANOL	C6H14O2
10	2-ETHOXYETHYL-ACETATE	C8H12O3
11	2-HEPTANONE	C7H14O
12	2-METHYL-1,3-PROPANEDIOL	C4H10O2
13	2-Methyl-2,4-pentandiol	C6H14O2
14	2-methylcyclohexanol	C7H14O
15	3-METHYL-1-BUTANOL	C5H12O
16	3-METHYL-2-BUTANOL	C5H12O
17	4-METHYL-2-PENTANOL	C6H14O
18	ACETYLACETONE	C5H8O2
19	CYCLOHEXANOL	C6H12O
20	Cyclohexanone	C6H10O
21	CYCLOHEXYL-ACETATE	C8H14O2
22	Cyclopentanol	C5H10O
23	Decane	C10H22
24	DIETHYL-OXALATE	C6H10O4
25	ETHYL-3-ETHOXYPROPIONATE	C7H14O3
26	ETHYLENE-GLYCOL-DIACETATE	C6H10O4
27	Hexadecane	C16H34
28	METHYL-ACETOACETATE	C5H8O3
29	TETRAHYDROFURFURYL ALCOHOL	C5H10O2

REFERENCES

Abildskov J. and O'Connell J.P., 2003. Predicting the solubilities of complex chemicals I. Solutes in different solvents. Industrial and Engineering Chemistry Research, 42, 5622-5634.

Achenie L.E.L., Gani R. and Venkatasubramanian V., 2003. Molecular design: theory and practice. Computer-Aided Chemical Engineering, 12. Elsevier, Amsterdam, the Netherlands.

Achenie, L. E. K., Churi, N., 1996. Novel Mathematical Programming Model for Computer Aided Molecular Design. Industrial & Engeneering Chemistry Research, 35(10), 3788-3794.

Achenie, L. E. K., Wang, Y., 2002. Computer aided solvent design for extractive fermentation. Fluid Phase Equilibria, 201(1), 1-18.

Bagajewicz M. J., 2007. On the role of macroeconomics, planning and finances in product design. AIChE Journal, 53, 3155-3170.

Bek-Pedersen, E., Gani, R., 2004. Design and synthesis of distillation systems using a driving-force-based approach. Chemical Engineering and Processing: Process Intensification, 43(3), 251-262.

Camarda, K. V., Maranas, C. D., 1999. Optimization in polymer design using connectivity indices. Industrial and Engineering Chemistry Research, 38, 1884-1892.

Cao, W., Knudsen K., Fredenslund A., Rasmussen, P., 1993. Simultaneous correlation of viscosity and vapour-liquid equilibrium data, Industrial Engineering Chemistry Research, 32, 2077-2087.

Ceriani, R., Gani, R., Liu, Y. A., 2013, Prediction of vapor pressure and heats of vaporization of edible oil/fat compounds by group contribution. Fluid Phase Equilibria, 337, 53-59.

Charpentier J.C., 2009. Perspective on multiscale methodology for product design and engineering. Computers and Chemical Engineering, 33, 936-946.

Charpentier, J. C., McKenna, T. F., 2004. Managing complex systems: some trends for future of chemical and process engineering. Chemical Engineering Science, 59(8-9), 1617-1640.

Chemdraw, Cambridgesoft, Inc., Cambridge, MA (2007) Retrieved from http://www.camsoft.com 4/1/2007

Chemmangattuvalappil, N. G., Solvason, C. C., Bommareddy, S., Eden, M. R., 2010. Molecular Signature Descriptors for Integrated Flowsheet and Molecular Design. Computer Aided Chemical Engineering, 28, 1267-1272.

Cheng, Y. S., Lam, K. W., Ng, K. M., Ko, R. K. M., Wibowo, C., 2009. An integrative approach to product development—A skin-care cream. Computers & Chemical Engineering, 33(5), 1097-1113.

Churi, N., Achenie, L. E. K., 1996. Novel Mathematical Programming Model for Computer Aided Molecular Design, Industrial & Engineering Chemistry Research, 35 (10), 3788-3794.

Churi, N., Achenie, L. E. K., 1997. The optimal design of refrigerant mixtures for a two-evaporator refrigeration system, Computers & Chemical Engineering, 21, s349-s354.

Cignitti, S., Zhang, L., Gani, R., 2015. Computer-aided framework for design of pure, mixed and blended products. Computer Aided Chemical Engineering, 37, 2093-2098.

Consantinou, L., Gani, R., 1994. New group contribution method for estimating properties of pure compounds. AIChE Journal, 40(10), 1697-1710.

Constantinou, L., Bagherpour, K., Gani, R., Klein, J. A., Wu., D. T., 1996. Computer-aided product design: problem formulations, methodology and applications. Computers & Chemical Engineering, 20(6/7), 685-702.

Conte, E., Gani, R., Malik, T. I., 2011. The virtual Product-Process Design laboratory to manage the complexity in the verification of formulated products. Fluid Phase Equilibria, 302, 294–304.

Costa, R., Moggridge, G. D., Saraiva, P. M., 2006. Chemical product engineering: an emerging paradigm within chemical engineering. AIChE Journal, 52(6), 1976-1986.

Cunico, L. P., Hukkerikar, A. S., Ceriani, R., Sarup, B., Gani, R., 2013, Molecular structure-based methods of property prediction in application to lipids: A review and refinement. Fluid Phase Equilibria, 357, 2-18.

Cussler, E. L., Wei, J., 2003. Chemical product engineering. AIChE Journal, 49(5), 1072-1075.

Davies, J.T., 1957. A quantitative kinetic theory of emulsion type, I. Physical chemistry of the emulsifying agent. Proceedings of the Gas/Liquid and Liquid/Liquid Interface Con-ference, 426-438.

Diaz-Tovar, C. A., Mustaffa, A. A., hukkerikar, A., Quaglia, A., Sin, G., Kontogeorgis, G., Sarup, B., Gani, R., 2011. Lipid Processing Technology: Building a Multilevel Modeling Network. Computer Aided Chemical Engineering, 29, 256-260.

Duvedi, A., Achenie, L. E. K., 1997. On the design of environmentally benign refrigerant mixtures: a mathematical programming approach, 21(8), 915-923.

Eden M.R. Jorgensen S.B., Gani R. and El-Halwagu M.M., 2004. A novel framework for simultaneous separation process and product design. Chemical Engineering and Processing, 43, 595-608.

Fedorova, M., Tolksdorf, G., Fillinger, S., Wozny, G., Sales-Cruz, M., Sin, G., Gani, R., 2015. Development of Computer Aided Modelling Templates for Model Re-use in Chemical and Biochemical Process and Product Design: Import and export of models. Computer Aided Chemical Engineering, 37, 953-958.

Fung, K. Y., Nakajima,S., Wibowo, C., Ng, K. M., 2006. A systematical Iterative Procedure for Determining Granulator Operating Conditions. AIChE Journal, 52, 3189-3202.

Gammon, J., 2004. Manual of Aviation Fuel Quality Control Procedure, ASTM Manual Series: MNL 5. West Conshohocken, Pa.: American Society for Testing and Materials.

Gani, R., 2004. Chemical product design: Challenges and opportunities. Computers & Chemical Engineering, 28, 2441–2457.

Gani, R., 2007. Case Studies in Chemical Product Design – Use of CAMD Techniques. Computer Aided Chemical Engineering, 23,435-458.

Gani, R., Bek-Pedersen, E., 2000, Simple Algorithm for Distillation Column Design. AIChE Journal, 46 (6), 1271.

Gani, R., Ng, K. M., 2015, Product design – Molecules, devices, functional products, and formulated products. Computers & Chemical Engineering, 81, 70-79.

Gani, R., Nielsen, B., Fredenslund, A., 1991. A Group Contribution Approach to Computer Aided Molecular Design. AIChE Journal, 37(9), 1318-1332.

Gani, R., Nielsen, B., Fredenslund, Aa, 1991. A Group Contribution Approach to Computer Aided Molecular Design. AIChE Journal, 37(9), 1318-1332.

Gani, R., Pistikopoulos, E. N., 2002. Property modelling and simulation for product and process design. Fluid Phase Equilibria, 194-197, 43-59.

Gani, R., Jiménez-González, C., Constable, D. J. C., 2005. Method for Selection of Solvents for Promotion of Organic Reactions, Computers and Chemical Engineering, 29, 1661-1676.

Garrett, T. K., 1994. Automotive Fuels and Fuel Systems, Vol. 2: Diesel, London: Pentech Press.

Gmehling, J., Bölts, R., 1996. Azeotropic Data for Binary and Ternary Systems at Moderate Pressures. Journal of Chemical & Engineering data, 41(2), 202-209.

Gross, J., Sadowski, G., 2001. Perturbed-Chain SAFT: An Equation of State Based on a Perturbation Theory for Chain Molecules. Industrial & Engineering Chemistry Research, 40, 1244-1260.

Gubbins, K. E., Quirke, N., 1996. Molecular Simulations and Industrial Applications. Amsterdam: Gordon and Breach.

Harper, P. M., 2000. A multi-phase, multi-level framework for computer aided molecular design. PhD dissertation, Technical University of Denmark, Denmark.

Harper, P.M., Gani, R., 2000. A multi-step and multi-level approach for computer aided molecular design. Computers & Chemical Engineering, 24, 677-683.

Heitzig, M., Gregson, C., Sin, G., Gani, R., 2011. Application of computer-aided multi-scale modelling framework - Aerosol case study. Computer Aided Chemical Engineering, 29, 16-20.

Hileman, J. I., Stration, R. W., 2014, Alternative jet fuel feasibility. Transport Policy, 34, 52-62.

Hill, M., 2009. Chemical Product Engineering – The third paradigm. Computer & Chemical Engineering, 33(5), 947-953.

Holland, J. H., 1975. Adoptation in natural and artificial systems. Ann Arbor, MI: University of Michigan Press.

Hostrup, M., Harper, P. M., Gani, R., 1999. Design of Environmentally Benign Processes: Integration of Solvent Design and Process Synthesis. Computers & Chemical Engineering, 23, 1394-1405.

Hostrup, M., Harper, P. M., Gani, R., 1999. Design of environmentally benign processes: integration of solvent design and separation process synthesis. Computer & Chemical Engineering, 23(10), 1395-1414.

Hukkerikar, A. S., Kalakul, S., Sarup, B., Young, D. M., Sin, G., Gani, R., 2012, Estimation of Environmental-Related Properties of Chemicals for Design of Sustainable Process: Development of Group Contribution+ (GC+) Property Models and Uncertainty Analysis. Journal of Chemical Information and Modeling, 52 (11), 2823-2839.

Hukkerikar, A. S., Sarup, B., Kate, A. T., Abildskov, J., Sin, G., Gani, R., 2012. Group-contribution+ (GC+) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. Fluid Phase Equilibria, 321, 25-43.

Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990 – 2005, 2007, EPA430-R-07-002, U.S. Environmental Protection Agency, Washington, United States of America.

Joback, K. G., Stephanopoulos, G., 1989. Designing molecules possessing desired physical property values. In Proceedings FOCAPD, 89, 363-387.

Joglekar G., Giridhar A. and Reklaitis G.V., 2014. A workflow modelling system for capturing data provenance. Computer & Chemical Engineering, 67, 148-158.

Kalakul, S., Ellegård, M. D., Abildskov, J., 2014. Solubility and Related Properties of Large Complex Chemicals – Part 3: Organic Solutes ranging from C6 to C40 in mixed solvents, vol. 15 of DECHEMA Chemistry Data Series. Frankfurt am Main: DECHEMA.

Kalakul, S., Hussian, R., Elbashir, N., Gani, R., 2015, VPPD Lab – The Chemical Product Simulator. Computer Aided Chemical Engineering, 37, 1415-1420.

Kalakul, S., Malakul, P., Siemanond, K., Gani, R., 2014. Integration of life cycle assessment software with tools for economic and sustainability analyses and process simulation for sustainable process design. Journal of Cleaner Production, 71, 98-109.

Karunanithi, A. T., Achenie, L., Gani, R., 2005, A new decomposition-based computer-aided molecular/mixture design methodology for the design of optimal solvents and solvent mixtures, Industrial and Engineering Chemistry Research, 44, 4785-4797.

Klatt, K., Marquardt, W., 2009. Perspectives for process systems engineering—Personal views from academia and industry. Computers & Chemical Engineering, 33(3), 536-550.

Koga, H., Itoh, A. Murayama, S., Suzue, S., Irikura, T., 1980. Structure-activity relationships of antibacterial 6,7 and 7,8-disubstituted 1-alkyl-1,4-hihydro-4-oxoquinoline-3-carboxylic acids. Journal of Medicinal Chemistry, 23, 1358-1363.

Kokossis, A., Marcoulaki, E., 1996. Stochastic optimization of complex reaction system. Computers & Chemical Engineering, 20, 1996, 231-236.

Korb, O., Stutzle, T., Exner, T. E., 2007. An ant colony optimization approach to flexible protein–ligand docking. Swarm Intelligence, 1(2), 115-134.

Li J., Mundhwa M., Tontiwachwuthikul P. and Henni A., 2007. Volumetric properties, viscosities, and refractive indices for aqueous 2-(methylamino)ethanol solutions. Journal of Chemical Engineering Data, 52, 560-565.

Liaw, H., Gerbaud, V., Li, Y., 2011. Prediction of miscible mixtures flash-point form UNIFAC group contribution methods. Fluid Phase Equilibria, 300(1-2), 70-82.

Lin, B.J.,Chen, L.J., 2004. Liquid-liquid equilibria for the ternary system water+tetra-decane+2-butoxyethanol. Fluid Phase Equilibria, 216, 13-20.

Magnussen, T., Rasmussen, P., Fredenslun, Aa., 1981. UNIFAC parameter table for prediction of liquid-liquid equilibria. Industrial and Engineering Process Design and Development, 20, 331-339.

Maranas, C. D., Floudas, C. A., 1997. Global optimization in generalized geometric programming. Computers & Chemical Engineering, 21(4), 351-369.

Marrero, J., Gani, R., 2001. Group-contribution based estimation of pure component properties. Fluid Phase Equilibria, 183-184, 183-208.

Mattei, M., Kontogeorgis, G. M., Gani, R., 2014. A comprehensive framework for surfactant selection and design for emulsion based chemical product design. Fluid Phase Equilibria, 362, 288–299.
Mitrofanov, I., Sansonetti, S., Abildskov, J., Gurkan, S., Gani, R., 2012. The Solvent Selection framework: solvents for organic synthesis, separation processes and ionic-organic synthesis. Computer Aided Chemical Engineering, 30, 762-766.

Modarresi, H., Conte, E., Abildskov, J., Gani, R., Crafts, P., 2008. Model-Based Calculation of Solid Solubility for Solvent Selections - A Review. Industrial & Engineering Chemistry Research, 47, 5234-5242.

Modi, A., Aumond, J. P., Mavrovouniotis, M., Stephanopoulos, G., 1996. Rapid plant-wide screening of solvents for batch processes. Computer & Engineering, 20, S375-S380.

Moggridge, G. D., Cussler, E. L., 2000. An Introduction to Chemical Product Design. Chemical Engineering Research & Design, 78, 5-11.

Moggridge, G. D., Cussler, E. L., 2003. Teaching chemical product design. Computer Aided Chemical Engineering, 15, 1460-1465.

Morales-Rodriguez, R., Singh, R., Cameron, I., Gani, R., 2011. Product and Process Modelling: A Case Study Approach. Elsevier. 363-432.

Morales-Rodriguez,R., Gani, R., 2009. Multiscale Modelling Framework for Chemical Product-Process Design. Computer Aided Chemical Engineering , 26, 495-500.

Muro-Suñé, N., Gani, R., Bell, G., Shirley, I., 2005, Model-based computer-aided design for controlled release of pesticides. Computer & Chemical Engineering, 30 (1), 28-41.

Ng K.M., Gani, R., Dam-hohansen,K., 2007, Chemical product design. Toward a perspective through case studies. Elsevier, Amsterdam, The Natherlands.

Ng, K. M., 2003. MOPSD: a framework linking business decision-making to product and process design. Computers & Chemical Engineering, 29(1), 51-56.

Ng, K. M., Gani, R., Dam-Johansen, K., 2007. Product Development – What to Make and How to Make. Computer Aided Chemical Engineering, 23, 473-489.

Nielsen, T.L., Abildskov, J., Harper, P.M., Papaeconomou, I., Gani, R., 2001. The CAPEC Database. Journal of Chemical Engineering Data, 46, 1041-1044.

Ourique, J. E., A. S. Telles, 1998. Computer-Aided Molecular Design with Simulated Annealing and Molecular Graphs. Computer & Chemical engineering, 22, S615.

Pretel, E., Lopez, P. A., Bottini, S. B., Brianole, E. A., 1994. Computer-aided design of solvents for separation processes. AIChE Journal, 40, 1349-1360.

Reid, R.C., J.M. Prausnitz, and B.E. Poling. 1987. The Properties of Gases and Liquids. United States: McGraw Hill Book Co.,New York, NY.
Salager, J.L., 1996. Quantifying the concept of physico-chemical formulation in surfactant-oil-water systems – State of the art. Progress in Colloid and Polymer Science, 100, 137-142.

Sayanarayana, K. C., Abildskov, J., Gani, R., 2009. Computer-aided polymer design using group contribution plus property models. Computers & Chemical Engineering, 33, 1004-1013.

Schneider, G., Hartenfeller, M., Reutlinger, M., Tanrikulu, Y., Proschak, E., Schneider, P., 2009. Voyages to the (un)known: adaptive design of bioactive compounds. Trends in Biotechnology, 27(1), 18-26.

Sinha M., Achenie, L. E. K., Gani, R., 2003, Blanket wash solvent blend design using interval analysis, Industrial and Engineering Chemistry Research, 42, 516-527.

Smith, B. V., Ierapepritou, M. G., 2011. Modeling and optimization of product design and portfolio management interface. Computers & Chemical Engineering, 35(11), 2579–2589.

Solvason, C. C., Chemmangattuvalappil, N. G., Eden, M. R., 2009. A systematic method for integrating product attributes within molecular synthesis. Computer & Chemical Engineering, 33, 977-991.

Suares, J. T., Torres-Marchal, C., Rasmussen, P., 1989. Prediction of surface tension of nonelectrolyte solutions. Chemical Engineering Science, 44 (3), 782-786.

Seider, W. D., Widagdo, S., Seader, J. D., Lewin, D. R., 2009. Perspectives on chemical product and process design. Computer & Chemical Engineering, 33(5), 930-935.

Seider, W. D., Seader, J. D., Lewin, D. R., Widagdo, S., 2008. Product and Process Design Principles: Synthesis, Analysis, and Design, 3rd Edition. Wiley, Now York, USA.

Seider, W.D., Seader J.D., Lewin D.R., Process Design Principles: Synthesis, Analysis, and Evaluation, 1999. John Wiley & Sons, Southern Gate, England.

Teixeira, M. A., Rodriguez, O., Rogrigues, S., Martins, I., Rodrigues, A. E., 2012. A case study of product engineering: Performance of microencapsulated perfumes on textile applications. AIChE Journal, 58, 1939-1950.

Venkatasubramanian, V., Chan, K., Caruthers, J. M., 1995. Generic algorithmic approach for computer-aided molecular design. Computer Aided Molecular Design, ACS Symposium, 589, 396-414.

Voncken, R. M., Broekhuis, A. A., Heeres, H. J., Jonker, H. G. H., 2004. The Many Facts of Product Technology. Chemical Engineering Research Design, 82(11), 1411-1424.

Wang, Y., Achenie, L. E. K., 2002. A hybrid global optimization approach for solvent design. Computers & Chemical Engineering, 26(10), 1415-1425.

Wang, Z.W., Li, G.Z., Zhang, X.Y., Wang, R.K., Lou, A.J., 2000. A quantitative structure-property relationship study for the prediction of critical micelle concentration of an-ionic surfactants. Colloid and Surfaces A: Physicochemical Engineering Aspects. 197, 37-45.

Young, D. M., Cebezas, H., 1999. Designing sustainable processes with simulation: the waste reduction (WAR) algorithm. Computer & Chemical Engineering, 23(10), 1477-1491.
Young, D., Scharp, R., Cebezas, H., 2000, The waste reduction (WAR) algorithm: enviromental impacts, energy consumption, and engineering economics. Waste Management, 20 (18), 605-615.

Yunus, N. A., Gernaey, K. V., Woodley, J. M., Gani, R., 2014. A systematic methodology for design of tailor-made blended products. Computers & Chemical Engineering, 66, 201-213.

Zhang, L., Cignitti, S., Gani, R., 2015. Generic mathematical programming formulation and solution for computer-aided molecular design. Computers & Chemical Engineering, 78, 79–84.

NOMENCLATURE

T_m	Normal melting point
T_b	Normal boiling point
T_c	Critical temperature
P_c	Critical pressure
V_c	Critical volume
ΔG_f^{298K}	Standard Gibbs free energy of formation at 298 K
ΔH_f^{298K}	Standard enthalpy of formation at 298 K
ω	Pitzer's ccentric factor
ΔH_{vap}^{298K}	Enthalpy of Vaporization at 298 K
ΔH_{vap}^{Tb}	Enthalpy of Vaporization at T_b
ΔH_f	Enthalpy of fusion
V_m^{298K}	Liquid molar volume at 298 K
Sur^{298K}	Liquid Surface Tension
δ_D^{298K}	Hansen Dispersive Solubility Parameter
δ_P^{298K}	Hansen Polar Solubility Parameter
δ_H^{298K}	Hansen Hydrogen-Bond Solubility Parameter
Log_{KOW}	Octanol/Water Partition Coefficient
Log_{WS}	Water Solubility Coefficient
pKa^{298K}	Acid Dissociation Constant
AiT	Auto Ignition Temperature
T_f	Flash point
$Visc_L^{298K}$	Liquid Viscosity at 298 K
$Therm.COND_L^{'}$	Liquid thermal conductivity
$-\log LC_{50}^{FM}$	Fathead Minnow 96-hr

$-\log LC_{50}^{DM}$	Daphnia Magna 48-hr
$-\log LD_{50}$	Oral Rat LD50
$logBCF$	Bio-concentration factor
$-$logPEL	Permissible exposure limit (OSHA-TWA)
PCO	Photochemical oxidation potential
GWP	Global warming potential
ODP	Ozone depletion potential
AP	Acidification potential
$-\log(EUAC)$	Emission to Urban Air (Carcinogenic)
$-\log(EUANon$	Emission to Urban Air (Non-Carcinogenic)
$-\log(ERAC)$	Emission to Rural Air (Carcinogenic)
$-\log(ERANon$	Emission to Rural Air (Non-Carcinogenic)
$-\log(EUAC)$	Emission to Urban Air (Carcinogenic)
$-\log(EUANon$	Emission to Urban Air (Non-Carcinogenic)
$-\log(ERAC)$	Emission to Rural Air (Carcinogenic)
$-\log(ERANon$	Emission to Rural Air (Non-Carcinogenic)
$-\log(EFWC)$	Emission to Fresh Water (Carcinogenic)
$-\log(EFWNor$	Emission to Fresh Water (Non-Carcinogenic)
$-\log(ESWC)$	Emission to Sea Water (Carcinogenic)
$-\log(ESWNon$	Emission to Sea Water (Non-Carcinogenic)
$-\log(EASC)$	Emission to Natural Soil (Carcinogenic)
$-\log(EASNon$	Emission to Agriculrural Soil (Non-Carcinogenic)
CP	Cloud Point
CMC	Critical Micelle Concentration
HLB	Hydrophilic-Lipophilic Balance
HHV	Higher heating value
CO_2E	CO_2 emission in combustion engine
Z_c	Critical Compressibility Factor
ΔS_{fus}	Entropy of Fusion

V_L^{Tb}	Liquid Volume at T_b
RD	Refractive Index
$olarRefractio$	Molar Refraction
DipolarMomen	Dipolar Moment
DielectricCons	Dielectric Constant
$HenryConst^{29}$	Henry Constant at 298 K
C	Cost
T_K	Krafft Temperature
$Diff.Coeff$	Diffusion coefficient at infinite dilution in water
ρ	Liquid Density
k	Liquid Thermal Conductivity
, P^{sat}	Vapor Pressure
ΔH_{vap}	Enthalpy of Vaporization
δ	Hildebrand Solubility Parameter
C_p^{ideal}	Ideal Gas Heat Capacity
C_p^{liquid}	Liquid Heat Capacity
σ	Liquid Surface Tension
η^{gas}	Vapor Viscosity
λ	Vapor Thermal Conductivity
V^{gas}	Vapor Volume
T90	Evaporation Time
η	Dynamic Liquid Viscosity
N_i, M_j, O_k	Occurrence number of group i
C_i	Contribution for the first-order group of type *i* with N_i occurrences
D_j	Contribution for the second-order group of type *j* with M_j occurrences
E_k	Contribution of the third-order group of type *k* with O_k occurrences
V_i	Volume of fuel type i

$HHVi$	Higher heating value of fuel type i
$C_{content,i}$	Carbon content coefficient of fuel type i
FO_i	Fraction oxidized of fuel type i
Ws	Water solubility
$\Delta H_{formation}$	Heat of formation
k	liquid thermal conductivity
Mw	molecular weight
C_p^{ideal}	ideal gas heat capacity
σ	surface tension
ŋ	gas viscosity; § is energy-potential parameter
T_r	reduced temperature
F_p^0	low pressure polar correction factor
F_Q^0	low pressure quantum correction factor
μ	dipole moment
μ_r	dimensionless dipole moment
λ	thermal conductivity
C_v	specific heat at constant volume
C_p	specific heat at constant pressure
$T90$	evaporation time
P_i	property of component *i*
γ_i	activity coefficient
V_{cm}	molar averages of the pure component critical volumes
Z_{RAm}	molar averages of the pure component critical compressibility factors
Z_{RAi}	particular constant for the Rackett equation for compound *i*
η_s	dynamic viscosity of the continuous phase
M	ratio between the dynamic viscosity of the dispersed phase and that of the continuous phase
ψ	volume fraction of the dispersed solvent phase
ECN	effective carbon number and it is related to the molecular

	structure of the organic solvent phase
T_0	initial boiling point
D_1	diffusion coefficient of AI in the polymer
D_0	constant pre-exponential factor
E	energy (per mole) that the molecule needs to overcome attractive forces which constrain it to its neighbors
R	gas constant
T_{gi}	glass transition temperature of compound i
w_i	weight fraction of component i
K_{1i}, K_{2i}	free volume parameter of compound i
$\widehat{V_t^*}$	critical hole free volume required for a jump
ξ	ratio of molar volumes for the solvent and the polymer jumping units
γ	overlap factor (between 0.5 and 1)
C	concentration of donor as function of time
$C_{d,initial}$	initial concentration of donor as function of time
$K_{m/d}$	partition coefficient of the AI between the donor and the polymer membrane
$K_{m/r}$	partition coefficient of the AI between the receiver and the polymer membrane
$V_{d,i}$	donor volume
V_r	receiver volume
A_i	surface area through which diffusion takes place
D	polymer solvent binary mutual diffusion coefficient
h	thickness of the microcapsule wall
t	time
$M_{r,i}\ (t)$	concentration of receiver as a function of time
$C'_{d,initial}$	modified initial concentration of donor as function of time
r	microcapsule radius
μ	mean distribution value
σ	standard deviation

Department of Chemical and Biochemical Engineering
Technical University of Denmark
Søltofts Plads, Building 229
2800 Kgs. Lyngby
Denmark

Phone: +45 45 25 28 00
Web: www.kt.dtu.dk/